EuroITV 2010

Chengyan Peng, Petri Vuorimaa, Pertti Näränen, Celia Quico,
Gunnar Harboe & Artur Lugmayr (Eds.)

Adunct Proceedings - EuroITV 2010

web.sharing.tv.content

Tampere, Finland, June 2010

EuroITV2010
9th June-11th June 2010, Tampere, Finland
www.euroitv2010.org

EMMi Lab. - www.tut.fi/emmi
Dept. of Business Inform. Management
Tampere University of Technology

NOKIA
CONNECTING PEOPLE

Microsoft Mediaroom

DEMOLAB

digitampere.fi

NoTube

YLE

amea
ambient * media * association
www.ambientmediaassociation.org

SIGCHI

acmsigmm

sigweb

acm

VTT

The City of Tampere

Tampere, Finland, 2010

Tampereen teknillinen yliopisto – Tampere University of Technology

Chengyan Peng, Petri Vuorimaa, Pertti Näränen, Celia Quico, Gunnar Harboe &
Artur Lugmayr (Eds.)

Adjunct Proceedings EuroITV 2010

web.sharing.tv.content

Tampereen teknillinen yliopisto
Tampere, Finland, June 2010

ISBN 978-952-15-2355-7

Preface

On behalf of EuroITV2010 organizing committee, it is our great pleasure to welcome you to Finland to the opening of the 8th European Conference on Interactive TV and Video (EuroITV2010) hosted by Tampere University of Technology in Tampere on 9-11 June, 2010. The conference is in-cooperation with ACM, SIGCHI, SIGWEB and SIGMM.

Distinguished guests are invited from academia and industrial professionals around the world to discuss the latest advances and research of media technology, HCI, media studies, and the content creation community. In this conference, the cutting edge and most recent topics concerning TV are presented and discussed from not only technological but also social and artistic points of view.

EuroITV has become increasingly competitive in terms of paper submission and acceptance rate. We received approx. 200 paper submissions relevant for the adjunct proceedings to this conference from 34 countries. Through peer reviewing process by program committee and experts, we accepted for the adjunct proceedings 26 posters, 5 doctoral consortium papers, 7 industrial cases, 8 demonstrations, 40 workshop papers, 2 panels, and 2 keynotes.

We have created a rich in content and high-quality technical program spanning three full days, covering four different tracks, representing the activities in TV research area. Presentations include keynote speeches by Jyri Huopaniemi from Nokia Research Center Tampere on "Real-time Context and Immersion Providing Uniquely New Mobile Experiences" and Marcos Gonzalez-Flower from Siemens IT Solutions and Services Ltd on "Future Challenges for Broadcasters and the Media Industry".

It is believed that this conference will surely give a great opportunity for all of you to meet together, exchange minds through the inspiring conversions, and think about the hints for research of new technology and business towards the future Television for the entertainment.

We would like to thank all members of the organizing committee for their dedication to the success of EuroITV2010 and timely review of the submissions. We would like to acknowledge Nokia, Microsoft, Hermia, Demola, digitampere, YLE, City of Tampere, and NoTube for their sponsorships – and especially VTT for their involvement in the editing process of the adjunct proceedings. We also thank our media partners, the Ambient Media Association (www.ambientmediaassociation.org), the Digital TV Group (DTG), iTV, digibusiness.eu, the Deutscher IPTV Verband, and informitv.

We would like to invite you to visit the city Tampere especially at this season of the year when nature is welcoming everybody and enjoy both the technical program and the opportunity to explore the city and the Nordic culture. Finally we sincerely hope that you can benefit attending this conference and enjoy staying in Finland. We hope your stay here will be rewarding and memorable.

Chengyan Peng, Petri Vuorimaa, Pertti Näränen, Celia Quico,
Gunnar Harboe & Artur Lugmayr

Table of Content

PREFACE ii
TABLE OF CONTENT iv
CALL FOR PAPERS xiv

Keynotes

Real-time Context and Immersion Providing Uniquely New Mobile Experiences 3
Jyri Huopaniemi, Nokia Research Centre Tampere, Finland, jyri.huopaniemi@nokia.com

Future Challenges for Broadcasters and the Media Industry 5
Marcos Gonzalez-Flower, Siemens IT Solutions and Services Ltd, UK, Flowe@Siemens.com

Panels

A Panel on How the Emergence of Internet TV is Impacting the User Experience of Traditional Pay TV 9
Michael J. Darnell, Microsoft TV, USA, mdarnell@microsoft.com

A Panel on the possible role of iTV in a multi-channel communication of Public Administration 10
Agnes Jenei, Corvinus University of Budapest, Hungary, agnes.jenei@uni-corvinus.hu

Doctoral Consortium

Understanding DTV Usages in the Era of Networked Communication: Integrating the Audience, the Media and the Public Policy's Perspectives 14
Vera Araújo, OberCom, Portugal, vera.araujo@obercom.pt

Converged Mobile Media: Evaluation of an Interactive User Experience 18
Alexandre Fleury, Aalborg University, Denmark, amf@es.aau.dk

Secondary-screen as a control device for TV content - Measuring user behavior 22
Francisco Javier Burón Fernández, Cordoba University, Spain, jburon@uco.es

Augmented local television and visual knowledge building Shared space learning objects and interactive services in a multiuser mixed reality Research Plan 26
John Grönvall, Arcada University of Applied Science, Finland, john.gronvall@arcada.fi

Upcoming Challenges for Streaming Media Content Distribution 30
Nafis Iskander, Tampere University of Technology, Finland, nafis.iskander@tut.fi
Artur Lugmayr, Tampere University of Technology, Finland, lartur@acm.org

Industrial Case Studies

Blending CGI with real 3D 37
Kasper Jordaens, VRT medialab, Belgium, kasper.jordaens@vrt.be

Redesign of VRT video archive search 38
Hendrik Dacquin, VRT medialab, Belgium, hendrik@studiomuscle.com

Web-Based, Interactive Dashboard for Live Pro Cycling TV Events 39
Hendrik Dacquin, VRT medialab, Belgium, hendrik@studiomuscle.com

Study on user engagement and interaction with overlay applications for ITV 40
Angélica Ñáñez, Barcelona Media - Center of Innovation, Spain, angelica.nanez@barcelonamedia.org
Raquel Navarro-Prieto, Barcelona Media - Center of Innovation, Spain, raquel.navarro@barcelonamedia.org

New Paradigm Models in Digital Media Business (IP TV and Mobile TV) 41
Zvezdan Vukanovic, University of Donja Gorica, Montenegro, zvezdanv@t-com.me

From Reach to Engage: Harnessing the Marketing Potential of Interactive TV 45
Ana C. Andrés del Valle, Accenture Technology Labs, France, ana.c.andresdelvalle@accenture.com

User experience of next generation content delivery: users are developing a new sense for quality 46
Marjan de Jonge, Oskar van Deventer, and Pieter Nooren, TNO Information and Communication Technology, The Netherlands, marjan.dejonge@tno.nl, oskar.vandeventer@tno.nl, and pieter.nooren@tno.nl

Tutorials

ESCape: Directions for the next wave of interactive television research 53
Konstantinos Chorianopoulos, Ionian University, Greece, choko@ionio.gr
Pablo Cesar, CWI, the Netherlands, p.s.cesar@cwi.nl

Tutorial: Designing and Evaluating the Sociability of Interactive Television and Online Video 55
David Geerts, K.U.Leuven, Belgium, david.geerts@soc.kuleuven.be

Tutorial: Experience prototyping layered ITV application 57
Rich Radka, Node – Innovation for Growth, Spain, richradka@gmail.com

Recommender Systems for Interactive TV 59
Paolo Cremonesi, Politecnico di Milano, Italy, paolo.cremonesi@polimi.it
Roberto Turrin, Neptuny, Italy, roberto.turrin@neptuny.com

Tutorial: Browser, Widgets and Applications: An Insight on State-Of-The-Art Technologies for the Creation of Service Frontends for Social and Interactive TV Experiences 61
Oliver Friedrich, Fraunhofer FOKUS, Germany, oliver.friedrich@fokus.fraunhofer.de

Demonstrations

Collaborative Annotation of Videos: watching and commenting (YouTube) videos.....66
Maria da Graça Pimentel, Universidade São Paulo, Brazil, mgp@icmc.usp.br
Roberto Faga Jr, Universidade São Paulo, Brazil, faga@icmc.usp.br
Vivian Motti, Universidade São Paulo, Brazil, vmotti@icmc.usp.br
Renan Cattelan, Univ.Federal de Uberlândia, Brazil, renan@facom.ufu.br
Cesar Teixeira, Univ.Federal de SCarlos, Brazil, cesar@dc.ufscar.br

The IITV Project, a video authoring tool based on recommendation 68
Toni Bibiloni, Balearic Island University, Spain , toni.bibiloni@uib.es
Pere Palmer, Balearic Island University, Spain, pere.palmer@uib.es
Miquel Mascaró, Balearic Island University, Spain, mascport@uib.es

Controlling Interactive Digital Television Applications by Brain-Computer Interface 71
Taciana Saad Rached, Laboratory Federal University of Campina, Brazil, taciana.rached@ee.ufcg.edu.br
Danilo Freire de S. Santos, Signove Tecnologia, Brazil, danilo.santos@signove.com
Angelo Perkusich, Federal University of Campina, Brazil, perkusic@dee.ufcg.edu.br
Hyggo Almeida, Federal University of Campina, Brazil, hyggo@dsc.ufcg.edu.br
Maurus Marques de Almeida Holanda, Federal University of Paraíba, Brazil, maurusholanda@hotmail.com

Text Input in Digital Television: a Component Prototype.....75
Diogo Pedrosa, Inst. de Ciências Matem. e de Computação, Brazil, diogopedrosa@uol.com.br
Didier A. Vega, Instituto de Ciências Matemáticas e de Computação, Brazil, davo@icmc.usp.br
Maria da Graça C. Pimentel, Inst. de Ciências Matem. e de Computação, Braizil, mgp@icmc.usp.br
Renata Pontin M. Fortes, Instituto de Ciências Matemáticas e de Computação, Brazil, renata@icmc.usp.br

Demonstration of the GAMECAST Interaction Environment.....79
Thomas Schmieder, University of Applied Sciences, Germany, Thomas@gamecast-tv.com
Benny Baumann, University of Applied Sciences, Germany, benbe@geshi.org
Christoph Schubert, University of Applied Sciences, Germany, Cschube5@hs-mittweida.de

SUMA & T-Maestro: Integrating personalized t-learning through standards 81
Silvia Bastos, Universidad de Vigo, Spain, sbastos@det.uvigo.es
Rebeca P. Diaz, Universidad de Vigo, Spain, rebeca@det.uvigo.es
Ana Fernández, Universidad de Vigo, Spain, avilas@det.uvigo.es
Gonzalo Jiménez, Universidad de Vigo, Spain, gonzalo@det.uvigo.es

Users in the loop: browsing archival video with advanced video retrieval technology 83
Johan Oomen, Netherlands Institute for Sound and Vision, The Netherlands, joomen@beeldengeluid.nl
Roeland Ordelman, University of Twente, The Netherlands, rordelman@beeldengeluid.nl
Cees G.M. Snoek, University of Amsterdam, The Netherlands, cgmsnoek@uva.nl
Bauke Freiburg, Video Dock, The Netherlands, bauke@videodock.com

Video Active and EUscreen Providing Audiovisual Content to Europeana 85
Johan Oomen, Netherlands Institute for Sound and Vision, The Netherlands, joomen@beeldengeluid.nl
Anna Christaki, National Technical University of Athens, Greece, achristaki@image.ntua.gr
Vassilis Tzouvaras, National Technical University of Athens, Greece, tzouvaras@image.ntua.gr

Posters

Mobile Video as an Instrument in a Social TV Service 92
Minna Suovirta, VTT Technical Research Centre of Finland, Finland, minna.suovirta@vtt.fi
Markus Ylikerälä, VTT Technical Research Centre of Finland, Finland, markus.ylikerala@vtt.fi
Chengyuan Peng, VTT Technical Research Centre of Finland, Finland, chengyuan.peng@vtt.fi

Understanding users of web-TV 96
Amela Karahasanović, University of Oslo, Norway, amela@sintef.no
Jan Heim, SINTEF ICT, Norway, Jan.Heim@sintef.no

Television Usability Tests with Conventional Remote Control and Wii Remote Using Prototypes 99
Christian P. Brackmann, Universidade Federal de Santa Maria, Brazil, christian@brackmann.com.br
Fabio R. Palamedi, Universidade Metodista de São Paulo, Brazil, fabio.palamedi@gmail.com

The need of user centred processes to do usable and accessible technology-design for DTV 103
Hans Persson, KTH, Sweden, hans.persson@iht.se
Francisco Fraile, Universidad Politécnica de Valencia, Sweden, francisco.fraile@itvarena.com
Dan Kopparhed, KTH, Sweden, dan.kopparhed@itvarena.com

Current Issues in Broadcasting from a Market Perspective 107
Artur Lugmayr, EMMi Lab, Finland, lartur@acm.org

Audience measurement system for digital television 111
Valdecir Becker, Escola Politécnica da USP, Brazil, valdecirbecker@gmail.com
Marcelo Knörich Zuffo, Escola Politécnica da USP, Brazil, mkzuffo@lsi.usp.br

What do you want to watch (again)? Video Navigation Using Recency Frames 115
Diogo Pedrosa, Universidade de São Paulo, Brazil, diogo@icmc.usp.br
Didier A. Vega, Universidade de São Paulo, Brazil, davo@icmc.usp.br
Maria da Graça C. Pimentel, Universidade de São Paulo, Brazil, mgp@icmc.usp.br

Experimental Characterisation of an Open Out-of-home MHP-based DVR Service 119
Philip Leroux, Ghent University-IBBT-INTEC, Belgium, philip.leroux@intec.ugent.be
Kris Vanhecke, Ghent University-IBBT-INTEC, Belgium, kris.vanhecke@intec.ugent.be
Klaas Roobroeck, Ghent University-IBBT-INTEC, Belgium, klaas.roobroeck@intec.ugent.be
Chris Develder, Ghent University-IBBT-INTEC, Belgium, chris.develder@intec.ugent.be
Filip De Turck, Ghent University-IBBT-INTEC, Belgium, filip.deturck@intec.ugent.be

Monitoring of Hypertensive Patients through Conventional Medical Devices Integrated to the Brazilian Digital TV 123
Vandermi J. Silva, Universidade Federal do Amazonas, Brazil, vandermi@ufam.edu.br
Vicente Ferreira de Lucena Júnior, Universidade Federal do Amazonas, Brazil, vicente@ufam.edu.br
Jomara Bindá, Universidade Estadual do Amazonas, Brazil, jomarabinda@gmail.com

Multi-user Interactive TV: the next step in personalization 127
Ray van Brandenburg M., TNO ICT, the Netherlands, ray.vanbrandenburg@tno.nl
Oskar van Deventer, TNO ICT, the Netherlands, oskar.vandeventer@tno.nl
Georgios Karagiannis, University of Twente, the Netherlands, karagian@cs.utwente.nl
Mike Schenk, TNO ICT, the Netherlands, mike.schenk@tno.nl

Design of an Architecture for Management of Context-Aware Narrowcasting 131
Klaas Roobroeck, Ghent University – IBBT, Belgium, klaas.roobroeck@intec.ugent.be
Philip Leroux, Ghent University – IBBT, Belgium, philip.leroux@intec.ugent.be
Tim Wauters, Ghent University – IBBT, Belgium, tim.wauters@intec.ugent.be
Filip De Turck, Ghent University – IBBT, Belgium filip.deturck@intec.ugent.be

Integrating Interactive TV with the Web 135

Cristina Carmona, CITIC, Spain, ccarmona@citic.es
José A. Rodríguez, CITIC, Spain, jarodriguez@citic.es
José A. Hurtado, CITIC, Spain, jahurtado@citic.es
J. R. Salinas, CITIC, Spain, jsalinas@citic.es

A study on location of context-aware services for Mobile Digital Television 139
André Valdestilhas, Santo André – São Paulo, Brazil, firmao@yahoo.com.br

Automatic video search using semantic concepts 142
Mats Sjöberg, Aalto University School of Science and Technology, Finland, mats.sjoberg@tkk.fi
Ville Viitaniemi, Aalto University School of Science and Technology, Finland, ville.viitaniemi@tkk.fi
Markus Koskela, Aalto University School of Science and Technology, Finland, markus.koskela@tkk.fi
Jorma Laaksonen, Aalto University School of Science and Technology, Finland, jorma.laaksonen@tkk.fi

Personalisation Technologies for IPTV channels. Mashup -TV project 146
Toni Bibiloni, University of Balearic Islands, Spain, toni.bibiloni@uib.es
Gabriel Fontanet, University of Balearic Islands, Spain, gfontanet@uib.es
Jorge Chao, University of Balearic Islands, Spain, jorge.chao@uib.es
Cristina Pérez, University of Balearic Islands, Spain, cristina.perez@uib.es

Media Content Management Platform 150
Francesc Pinyol Margalef, UNIVERSITAT RAMON LLULL, Spain, fpinyol@salle.url.edu
Marc Rovira Vall, UNIVERSITAT RAMON LLULL, Spain, marcr@salle.url.edu
Jordi Borrell, UNIVERSITAT RAMON LLULL, Spain, jordibs@salle.url.edu
Gabriel Fernandez Ubiergo, UNIVERSITAT RAMON LLULL, Spain, gabrielf@salle.url.edu

RAUDOS: Interactive Multiplatform Audiovisual Delivery 154
Eduardo Burgoa, AIDO, Spain, eburgoa@aido.es
Marc Rovira, LA SALLE, Spain, marcr@salle.url.edu
Fernando García, ANDAGO, Spain, fernando.garcía@andago.es

DVB-PCF to Flash Lite Transcoder 158
Ramon Martín de Pozuelo, UNIVERSITAT RAMON LLULL, Spain, ramonmdp@salle.url.edu
Francesc Enrich, UNIVERSITAT RAMON LLULL, Spain, frane@salle.url.edu
Gabriel Fernández, UNIVERSITAT RAMON LLULL, Spain, gabrielf@salle.url.edu

Using video and social networks to improve the access of daycare centers to museums: one case study 162
Diana de Sousa, Aalto University School of Art and Design, Finland, dspsousa@gmail.com
Mariana Salgado, Aalto University School of Art and Design, Finland, mariana.salgado@gmail.com

Tele Graph: Information Enriched Interactive WEBTV 166
Mehmet Özer Metin, Hacettepe Üniversitesi Beytepe,Turkey, ozer.metin@anelarge.com
Siyamed Sınır, Hacettepe Üniversitesi Beytepe, Turkey, siyamed.sinir@anelarge.com

Interactive TV-centric Health Care Services in Smart Homes 170
Pau Arce, Universidad Politécnica de Valencia, Spain, paarvi@iteam.upv.es
Francisco Fraile, Universidad Politécnica de Valencia, Spain, ffraile@iteam.upv.es
Juan Carlos Guerri, Universidad Politécnica de Valencia, Spain, jcguerri@dcom.upv.es

Derrick Metadata Platform An Internet Platform for Television Metadata 174
Alberto Attilio Reggiori, Derrick Media Ltd, UK, alberto@derrickmedia.com

A Software Product Line Requirements and Reference Architecture for TV Navigation System 178
Antonio Ferreira dos Santos Jr., Universidade Federal do Amazonas, Brazil, ajunior.analista@gmail.com
Vicente Ferreira de Lucena Jr., Universidade Federal do Amazonas, Brazil, vicente@ufam.edu.br

Evaluation of a TV Programs Recommendation using the EPG and Viewer's Log Data 182
Kazuki Ikawa, University of Tokyo, Japan, kazuki@parkcity.ne.jp
Tomohiro Fukuhara, National Institute of Advanced Industrial Science and Technology, Japan, tomohiro.fukuhara@aist.go.jp
Hideki Fujii, University of Tokyo, Japan, fujii@save.sys.t.u-tokyo.ac.jp
Hideaki Takeda, National Institute of Informatics, Japan, takeda@nii.ac.jp

YouToRemix : the online audiovisual quotting live remix application project 186
Milena Szafir, Manifesto21.TV, BRAZIL, milena@manifesto21.com.br

Mobile TV: Cultural Based Design Evaluation 193
Menandros Apostolidis, ARX NET S.A., Greece, chriskel@otenet.gr
Maaruf Ali, Oxford Brookes University, UK, mali@brookes.ac.uk

Workshop 1: Interactive Digital TV in Emergence Countries (BRIC)

Development of an Academic Laboratory of Interactive Digital TV in Amazon, Objectives and Main Projects 200
Vicente Ferreira de Lucena Jr, University of Amazonas, Brazil, vicente@ufam.edu.br
Waldir Sabino da Silva Jr, University of Amazonas, Brazil, waldirjr@ufam.edu.br
Celso Barbosa Carvalho, University of Amazonas, Brazil, celso@gta.ufrj.br
José Pinheiro de Queiroz Neto, Federal Institute of Education, Brazil, pinheiro@ifam.edu.br

China's IPTV Industry, Status quo and Trend 206
Zhen Yu GU, Shanghai Jiao Tong University, China, zygu@sjtu.edu.cn
Lin GUI, Shanghai Jiao Tong University, China, Guilin@ sjtu.edu.cn

Interactive Digital TV Services in India 209
Rajendra Akerkar, Western Norway Research Institute, Norway, rak@vestforsk.no

Ginga Edutainer: A Framework for the Development of T-Learning Applications on Ginga – The Brazilian Digital Television 212
Newton Nyamasege Marube, Federal University of Rio Grande, Brazil, nyamasege@gmail.com
Daniel Campelo Tavares, Federal University of Rio Grande, Brazil, prof.danieltavares}@gmail.com

Interactive Virtual Worlds in Brazilian Digital Television 217
Daniel Campelo Tavares, Federal University of Rio Grande, Brazil, prof.danieltavares,@gmail.com,
Newton Nyamasege Marube, Federal University of Rio Grande, Brazil, nyamasege@gmail.com
Silvia Silva da Costa Botelho, Federal University of Rio Grande, Brazil, silviacb.botelho@gmail.com
Nelson Lopes Duarte Filho, Federal University of Rio Grande, Brazil, dmtnldf@furg.br

An Application Model for Digital Television in e-Learning 221
Paloma Maria Santos, Federal University of Santa Catarina, Brazil, pmariasantos@yahoo.com.br
Marcus de Melo Braga, Federal University of Santa Catarina, Brazil, marcus@egc.ufsc.br
Marcus Vinícius Anátocles da Silva Ferreira, Federal University of Santa Catarina, Brazil, marcus.ferreira@unisul.br
Fernando José Spanhol, Federal University of Santa Catarina, Brazil, spanhol@led.ufsc.br

Conceptual Models for T-Commerce in Brazil 225
Bruno Cavaler Ghisi, Federal University os Santa Catarina, Brazil, bcg@certi.org.br
Guilherme Figueiredo Lopes, Brava ECM, Brazil, guilherme.lopes@bravaecm.com.br
Frank Siqueira, Federal University os Santa Catarina, Brazil, frank@inf.ufsc.br

An Alternative for the Interaction with Digital TV 231
Fábio de Jesus Lima Gomes, UFRGS, Brazil, fjlgomes@inf.ufrgs.br
José Valdeni de Lima, UFRGS, Brazil, valdeni@inf.ufrgs.br
Rosane Aragon de Nevado, UFRGS, Brazil, rosane@edu.ufrgs.br

The Importance of Digital TV for Countries in Development: a case study of Brazil 237
Hildegard Paulino Barbosa, MUS3-UFPB, hildegardpaulino@gmail.com
Julio Cesar Ferreira da Silva, LAVID-UFPB, Cidade Universitária, Brazil, julio@lavid.ufpb.br
Ed Porto Bezerra, UFPB, Brazil, edporto@di.ufpb.br

An Analysis of iDTV Appropriation in Developing Countries from the Users' Behavior Involved in Community-Oriented Content Creation Process 242
Elizabeth S. Furtado, UNIFOR/UECE, Brazil,Elizabet@unifor.br
Vasco Furtado, UNIFOR/SETICE, Brazil, vasco@unifor.br
Edson Santiago, UNIFOR, Brazil, edsonsantiago@uol.com.br

The Digital TV as another solution to educate in isolated areas in the Amazon State, Brazil 246
Lílian Simão Oliveira, Universidade de São Paulo, Brazil, lilianso@icmc.usp.br
Maria da Graça Pimentel, Universidade de São Paulo, Brazil, mgp@usp.br
José Pinheiro Queiroz-Neto, Instituto Federal Amazonas, Brazil, pinheiro@ifam.edu.br

Workshop 2: Methods for User Studies of Interactive (TV) Technologies

SocialMedia Experience and Design Lab: Using Advantages from Different LivingLab-Approaches 254
Jan Hess, University of Siegen, Germany, jan.hess@uni-siegen.de
Corinna Ogonowski, University of Siegen, Germany, corinna.ogonowski@uni-siegen.de

Using a Method Mix to Overcome Limitations in Studies of Technology Usage in the Home 256
Michael Pirker, Université Paul Sabatier Toulouse, France, Michael.Pirker@irit.fr
Mahmoud Abdellatif, Université Paul Sabatier Toulouse, France, Mahmoud.Abdellatif@irit.fr
Regina Bernhaupt, ruwido user experience research, Austria, Regina.Bernhaupt@ruwido.com

Methods for user studies of interactive TV in a commercial context 258
Paul Marrow, BT Innovate & Design, UK, paul.marrow@bt.com
Sue Hessey, BT Innovate & Design, UK, sue.hessey@bt.com
Ian Matthews, BT Innovate & Design, UK, ian.r2.matthews@bt.com

Concept and Usability Testing Methodologies for Interactive TV 260
D.J. Wheatley, Motorola Applied Research Center, USA, David.j.Wheatley@motorola.com
S. Basapur, Motorola Applied Research Center, USA, sbasapur@motorola.com

A Pragmatic Approach to Testing Issues in a Mobile Platform that Does Not Yet Exist: The CAMMP Story 262
Alexandre Fleury, Aalborg University, Denmark, amf@es.aau.dk
Jakob Schou Pedersen, Aalborg University, Denmark, jsp@es.aau.dk
Lars Bo Larsen, Aalborg University, Denmark, lbl@es.aau.dk

Open Profiling of Quality as a mixed method approach to study multimodal experienced quality ... 264
Dominik Strohmeier, Ilmenau University of Technology, Germany, dominik.strohmeier@tu-ilmenau.de

Triangulating UX Methods for Targeting End-Users' Subjective Experiences of Media Content 266
Annika Wiklund-Engblom, Åbo Akademi University, Finland, annika.wiklund-engblom@abo.fi
Anette Bengs, Åbo Akademi University, Finland, anette.bengs@abo.fi
Susanne Sperring, Åbo Akademi University, Finland, susanne.sperring@abo.fi

Using Multiple iTV Displays for the Simultaneous Comparative Evaluation of Parallel Prototypes ..268
Pierre-Alexandre Lapointe, Yu Centrik, Canada, palapointe@yucentrik.ca
Jay Vidyarthi, Yu Centrik, Canada, jay@yucentrik.ca
Joelle Stemp, Yu Centrik, Canada, jstemp@yucentrik.ca

Workshop 3: New Dimensions in the Assessment and Support of Quality of Experience (QoE) for Multimedia Applications)

Controlled vs. Uncontrolled Degradations of QoE – The Provisioning-Delivery Hysteresis in Case of Video 273
Thomas Zinner, University of Wuerzburg, Germany, fzinner@informatik.uniwuerzburg.de
Tobias Hossfeld, University of Wuerzburg, Germany, hossfeldg@informatik.uniwuerzburg.de
Tahir Nawaz Minhas, Blekinge Institute of Technology, Sweden, ftnm@bth.se
Markus Fiedler, Blekinge Institute of Technology, Sweden, mfig@bth.se

Subjective Quality Evaluation of H.264 High-Definition Video Coding versus Spatial Up-Scaling and Interlacing 276
Yohann Pitrey, École Polytechnique de l'Université de NANTES, France, yohann.pitre@univ{tes.fr
Marcus Barkowsky, École Polytechnique de l'Université de NANTES, France, marcus.barkowsky@univ{tes.fr
Patrick Le Callet, École Polytechnique de l'Université de NANTES, France, patrick.lecallet@univ{tes.fr
Romuald Pépion, École Polytechnique de l'Université de NANTES, France, romuald.pepion@univ{tes.fr

YoMo: A YouTube Application Comfort Monitoring Tool 280
Barbara Staehle, University of Würzburg, Germany, bstaehle@informatik.uni-wuerzburg.de
Matthias Hirth, University of Würzburg, Germany, hirth@informatik.uni-wuerzburg.de
Rastin Pries, University of Würzburg, Germany, pries@informatik.uni-wuerzburg.de
Florian Wamser, University of Würzburg, Germany, wamser@informatik.uni-wuerzburg.de
Dirk Staehle, University of Würzburg, Germany, dstaehle@informatik.uni-wuerzburg.de

Using metadata in video quality assessment based on the structural similarity (SSIM) index metric 283
Tiago Gonçalves, Instituto Universitário de Lisboa, Portugal, a24548@iscte.pt
Rui J. Lopes, Instituto Universitário de Lisboa, Portugal, Rui.Lopes@iscte.pt
Paulo Nunes, Instituto Universitário de Lisboa, Portugal, Paulo.Nunes@iscte.pt

Test persons for subjective video quality testing: Experts or non-experts? 286
Matej Nezveda, University of Vienna, Austria, matej.nezveda@univie.ac.at
Shelley Buchinger, University of Vienna, Austria, shelley.buchinger@univie.ac.at
Werner Robitza, University of Vienna, Austria, werner.robitza@univie.ac.at
Ewald Hotop, University of Vienna, Austria, ewald.hotop@univie.ac.at
Patrik Hummelbrunner, University of Vienna, Austria, patrik.hummelbrunner@univie.ac.at
Helmut Hlavacs, University of Vienna, Austria, helmut.hlavacs@univie.ac.at
Vittorio Baroncini, Fondazione Ugo Bordoni, Italy, vittorio@fub.it
Cristina Delogu, Fondazione Ugo Bordoni Italy, cristina@fub.it

Towards User-driven Adaptation of H.264/SVC Streams 289
Jordi Ortiz Murillo, University of Murcia, Spain, jordi.ortiz@um.es
Michael Ransburg, Klagenfurt University, Austria, mransbur@itec.uni-klu.ac.at
Eduardo Martínez Graciá, University of Murcia, Spain, edumart@um.es
Michael Sablatschan, Klagenfurt University, Austria, msablats@itec.uni-klu.ac.at
Antonio F. Gómez Skarmeta, University of Murcia, Spain, skarmeta@um.es
Hermann Hellwagner, Klagenfurt University, Austria, hellwagn@itec.uni-klu.ac.at

Enabling distributed fault diagnosis in multicast-based IPTV delivery 293
Jorma Kilpi, VTT, Finland,Jorma.Kilpi@vtt.fi

Spatio-Temporal Quality of Experience Trade-offs for Mobile Imaging Applications 296
Muhammad Imran Iqbal and Hans-Jürgen Zepernick, Blekinge Institute of Technology, Sweden, mii@bth.se
Markus Fiedler and Junaid Shaikh, Blekinge, Institute of Technology, Sweden, mfi@bth.se

Workshop 4: Future Television: Integrating the Social and Semantic Web

Linking TV And Web Using Semantics, A NoTube Application 302
Balthasar Schopman, Vrije Universiteit, The Netherlands, schopman@cs.vu.nl
Davide Palmisano, Fondazione Bruno Kessler, Italy, davide@asemantics.com
Ronald Siebes, Vrije Universiteit, The Netherlands, rm.siebes@few.vu.nl
Chris van Aart, Vrije Universiteit, The Netherlands, chris@vanaart.com
Véronique Malaisé, Vrije Universiteit, The Netherlands, vmalaise@few.vu.nl
Michele Minno, Pro-netics, Italy, michele.minno@asemantics.com

Semantic TV-resources brokering towards future television 310
Hong Qing Yu, The Open University, UK, h.q.yu@open.ac.uk
Stefan Dietze, The Open University, UK, s.dietze@open.ac.uk
Neil Benn, The Open University, UK, n.j.l.benn@open.ac.uk

Integrating social networks in an IPTV recommender system 315
Cerdeira Frederico, University of Aveiro Portugal, frederico.cerdeira@ua.pt
Almeida Pedro, University of Aveiro, Portugal, almeida@ua.pt
Abreu Jorge, University of Aveiro, Portugal, jfa@ua.pt
Almeida Pedro L., University of Aveiro, Portugal, pedroalmeida@ua.pt
Cardoso Bernardo, PT Inovação, Portugal, bernardo@ptinovacao.pt

Digital Video Interaction to support frontal lecturing in multiple platforms of internet based communication 319
Francesco Epifania, Politecnico di Milano, Italy, Epifania@elet.polimi.it
Giorgio Valle, Università degli Studi di Milano, Italy, Giorgio.valle@unimi.it

Workshop 5: Bridging among People, Places and Devices by Integrated, Ambient and Playful Social Media Approaches

An Interactive Couch Table to support TV-centric Social Interactions between Households 330
Lin Wan, International Institute for Socio-Informatics, Germany
Marcel Tweer, International Institute for Socio-Informatics, Germany

Designing a Social ITV Field Trial for real customers 333
Abreu Jorge, University of Aveiro, Portugal, jfa@ua.pt
Almeida Pedro, University of Aveiro, Portugal, almeida@ua.pt
Figueiredo Cátia, University of Aveiro, Portugal, cat@ua.pt
Pinto Ricardo, University of Aveiro, Portugal, ricardopinto@ua.pt
Teixeira Vera, University of Aveiro, Portugal, vmmt@ua.pt

Designing a large social display for an old people's home 335
Claudia Müller, University of Siegen, Germany, claudia.mueller@uni-siegen.de
Cornelius Neufeldt, University of Siegen, Germany, cornelius.neufeldt@uni-siegen.de
Lars Schröer, University of Siegen, Germany, lars.schroeer@uni-siegen.de

Constraints and simplification for a better mobile video nnotation and content customization process 339
Robert Knauf, Chemnitz University of Technology, Germany, robert.knauf@cs.tu-chemnitz.de
Arne Berger, Chemnitz University of Technology, Germany, arne.berger@cs.tu-chemnitz.de
Maximilian Eibl, Chemnitz University of Technology, Germany, maximilian.eibl@cs.tu-chemnitz.de

TV as an interactive medium to reflecting Energy consumption in daily life 343
Tobias Schwartz, Fraunhofer Institute for Applied Information Technology, Germany, tobias.schwartz@fit.fraunhofer.de
Johanna Meurer, Fraunhofer Institute for Applied Information Technology, Germany, johanna.meurer@fit.fraunhofer.de
Matthias Beztz, Fraunhofer Institute for Applied Information Technology, Germany, matthias.betz@fit.fraunhofer.de

Enable Users to choose from different Channels – Using Multichannel Technology to enrich the TV Experience 345
Sebastian Groß, Pixelpark AG, Germany, sebastian.gross@pixelpark.com
Sabrina Josellis, Pixelpark AG, Germany, sabrina.josellis@pixelpark.com
Roland Westermaier, Pixelpark AG, Germany, roland.westermaier@pixelpark.com

Understanding Mobile Social Networks 347
Benjamin Mangold, University of Siegen, Germany, mangold@ifm.uni-siegen.de

Ambient-Aware Service Infrastructure for Home IT Environments 350
Benedikt Ley, University of Siegen, Germany, benedikt.ley@uni-siegen.de
Martin Stein, University of Siegen, Germany, martin.stein@uni-siegen.de

---

8th European Conference on Interactive TV and Video EuroITV2010 "web.sharing.tv.content"

June 9th-11th, Tampere, Finland, http://www.euroitv2010.org/

Call for Papers, Targeted Workshops, Workshops, Tutorials, and Industrial Exhibitions

New in EuroITV:
- 4 tracks: interactive content & arts, HCI, technology, and media studies
- video and TV in the age of Web 3.0.
- EuroITV competition grand challenge
- targeted and themed workshops
- email list: on https://listmail.tut.fi/mailman/listinfo/euroitv2010
- proceedings published within the ACM Digital Library and special issues
- first deadlines approaching!

EuroITV is the leading international conference for media related to video and television. EuroITV is attended by academia and professionals from all over the world to discuss latest advances and research of media technology, HCI, media studies, and the content creation community. In addition to previous years' topics of the conference, EuroITV 2010 offers new opportunities this year:

* INTERACTIVE CONTENT & ARTS TRACK
EuroITV 2010 extends its tracks with a new track to explicitly attract the creative media production community dealing with interactive video, 3D cinema, TV program formats, film production, and consumers as content creators and contributors.

* VIDEO IN THE AGE OF WEB3.0.
EuroITV 2010 faces the challenges of TV as changing environment and explicitly invites contributions dealing with user-generated content, video in social networking, mashup networks, IPTV, cross-media, broadcasting everywhere, 3D cinema/3D IPTV/3D content, and video.

* EuroITV COMPETITION GRAND CHALLENGE
EuroITV 2010 also offers a competition opened for artistic content creators, application developers, and service designers to get awarded by EuroITV.

* TARGETED WORKSHOPS
We are specifically calling for targeted workshops around the themes mentioned below. In difference to general workshops, targeted workshops shall help us to explore current topics of interest in the community. Targeted workshop contributions will be published within the main proceedings, and the ACM Online Digital Library.

. Everything in 3D
(e.g. production, interactive content, digital games, 3D-TV, 3D-IPTV, 3D cinema, standards, 3D distribution);
. TV and video in emerging countries
(e.g. policies, technologies, state-of-the-art, middleware, content models, interactivity in countries such as China, South America, India, Russia);
. Content & art for/in TV and video
(e.g. interactive art works, tools, production methods, artistic designs, collaborative content, art forms, interactive designs, art collections, interaction methods);
. Visual communication for video, film, and TV
(e.g. visual language, pervasive communication, semiotics, storyline);
. Media management and economics
(e.g. patterns of use and engagement, business models, advertising, strategies, sustainability, policy, model of audiences, public service broadcasting, new media and traditional media);
. Social media, ambient/ubiquitous media, and digital games in the age of Web 3.0
(e.g. context awareness, next generation user interfaces, personalization, ambient human-computer interaction, hardware/software, interactive game design for video environments, serious gaming, persuasive games, social media tools, social networks mashups, community design);
. Emerging topics around the theme "web.sharing.tv.content"
(e.g. visionary - pointing-to-the future topics and issues around the conference theme);

IMPORTANT DATES

November 30th, 2009: Targeted Workshop Proposals
November 30th, 2009: Tutorials and Workshop Proposals
December 23rd, 2009: Full Paper Submissions
February 07th, 2010: Short Papers, Posters, Doctoral Consortium, Demos
March 21st, 2010 : Industry Exhibitions
April 1st, 2010 : EuroITV Competition Grand Challenge

More information about other sessions, e.g. for business oriented work, will follow soon.
More details can be found on: http://www.euroitv2010.org/

CONFERENCE TRACKS

This year's conference consists of four tracks, each with their own program committee and a separate track chair, coordinated by an overall Program Chair. Full papers, short papers and posters have to be submitted in one of these tracks, to optimize the review process. If not sure, choose the track which is closest to the main topic of your submission, and of which you feel that community is most appropriate to review your work.

Track 1: Interactive Content & Arts
Track 2: Human-Computer Interaction
Track 3: Media, Social and Economic Studies
Track 4: Systems and Enabling Technologies

Papers are solicited from, but not limited to the following topics:

. Storytelling, role playing, and language for video & TV
. Virtual worlds, mixed reality, and bringing video into the world
. Digital gaming, ambient media, social media, and locative media
. Artistic and aesthetic uses of video & TV technology
. Beyond the home context, extended home, Mobile TV
. Ambient intelligence, ambient media environments
. Social TV, sociability, usability and user experience
. Digital content production, HDTV and digital cinema
. Asset management, metadata and content enrichment
. Entertainment computing, games, betting, game shows
. Broadband, IPTV, 3DTV and VR systems
. Audience research, television studies, ethnography, user studies
. New advertising and revenue models for television
. Accessibility, universal access, multimodal interaction
. Business models, media management, media economics, t-commerce, t-learning
. Web2.0, social media, community television, user-generated content
. Communication services, video conferencing, messaging
. Content management, digital rights management
. Interactive storytelling, interactive advertising
. Electronic program guide, video search, video navigation
. Enhanced TV (news, weather, sports)
. Changes in technical requirements and infrastructures (ubiquitous and mobile)
. Standards (TV-Anytime, MPEG-4, MPEG-7, SMIL)
. Multimedia, graphics, broadcast and video technology
. Personalization, user modelling, intelligent user interfaces
. Ethical, regulatory and policy issues
. Everyday life practices by family, elderly, youngsters and children
. Digital divide and e-inclusion issues
. Methods for digital television research and design

Track 1: Interactive Content & Arts Track
. Studies of interactive, collaborative content & video
. Storytelling
. Digital gaming
. Role playing
. Virtual worlds, mixed reality, and bringing video into the world
. Distributed and networked experience design
. Immersion
. Social media in storytelling

- Ambient media, locative media, cross media formats, and social media
- Artistic and aesthetic uses of interactivity
- Viewer, player, user, and collaborative co-creator
- Towards a new form language for video & television
- Social media utilizing online video as an art form
- Video and TV genres on interactive platforms and in virtual worlds
- Video and TV as alternate reality games

Track 2: Human-Computer Interaction
- User-centered design
- Development and evaluation of iTV systems and services
- Interaction technologies for interactive TV and cross-media technologies
- Studies on iTV and media usage
- Design and evaluation methods for the area of interactive TV
- Case studies of successful/unsuccessful iTV systems and deployments of usability and user experience
- Accessibility of iTV services
- Cross-over media concepts and implementations
- Investigations on iTV acceptance and related concepts (trust, privacy, security).

Track 3: Media, Social and Economic Studies
- Television practices and experiences related to privacy, trust and security
- Direct marketing and profiling techniques for interactive TV: personalization and user modelling
- Measurability, interactivity and addressability of ITV use and prosumption
- Business models, media management, media economics, t-commerce, t-learning
- Interactive Digital Television and internet-of-things applications
- Audience research, television studies, ethnography, user studies
- New advertising and revenue models for television
- User creativity on/via television: Web2.0, social media, community television, user-generated content
- Ethical, regulatory and policy issues
- Everyday life practices by family, elderly, youngsters and children
- Digital divide and e-inclusion issues
- Methods for digital television research and design
- From mass media to mass self-communication
- Transitions in broadcasting industry and policy related to digital convergence and experience economy

Track 4: Systems and Enabling Technologies
- Mobile TV
- Ambient intelligence
- Social TV
- Digital content production
- HDTV and digital cinema
- Entertainment computing
- Interactive services (games, betting, game shows)
- Broadband, IPTV, 3DTV and VR systems
- Accessibility, universal access, multimodal interaction
- Web 2.0
- Internet TV (P2P TV, Web-based Interactive TV)
- Enhanced TV (news, weather, sports)
- Multimedia communication services
- Video technology (video conferencing, broadcast, video search)
- Standards (TV-Anytime, MPEG-4, MPEG-7, SMIL)

Paper submissions will be peer-reviewed. The main proceedings will be published by ACM, and be made available in the ACM Digital Library (http://portal.acm.org/). Extended versions of selected papers will be considered for special issues in a journals or books.

CONFERENCE ORGANIZING COMMITTEE

Conference Chairs
- Artur Lugmayr, EMMi Lab., Tampere Univ. of Technology (TUT), FINLAND
- Petri Vuorimaa, Helsinki Univ. of Technology (HUT), FINLAND

. Pertti Naranen, Tampere Univ. of Applied Sciences (TAMK), FINLAND

Program Chairs
. Celia Quico, Universidade Lusófona, Lisbon, PORTUGAL
. Gunnar Harboe, Motorola, U.S.A

Publicity Chairs
. Erika Reponen, Nokia Research Centre Tampere, FINLAN

EuroITV Competition Grand Challenge
o Susanne Sperring, MediaCity, FINLAND
o Robert Strezbowski, Beuth Hochschlue fuer Technik Berlin, GERMANY
o Milena Szafir, BRAZIL

Sponsor Chair
. Pauliina Tuomi, TUT, FINLAND

Tutorials Chair
. Marianna Obrist, HCI & Usability Unit, ICT&S, University of Salzburg, AUSTRIA
. David Geerts, CUO, IBBT / K.U. Leuven, BELGIUM

Doctoral Consortium Chair
. Hendrik Knoche, University College London, UK
. Judith Masthoff, University of Aberdeen, UK

Industrial Exhibition and Demonstration Chairs
. Jussi Pekka Koskiranta, YLE, FINLAND
. Gene Chipman, StreamSage, U.S.A.
. Timo Argillander, Digital Media Finland Oy, FINLAND

Short Papers & Posters Chairs
. Jens Jensen, Aalborg University, DENMARK
. Lyn Pemberton, Brighton University, UK

Workshop Chairs
. Pablo Cesar, CWI, THE NETHERLANDS
. Jari Multisilta, Tampere Univ. of Technology (TUT), Pori, FINLAND

Track chairs
HCI
. Regina Bernhaupt, HCI & Usability Unit, ICT&S, University of Salzburg (A)
. Oliver Friedrich, Fraunhofer Institute, FOKUS, GERMANY

Media Studies
. Jo Pierson, IBBT-SMIT, VUB (B)
. Jaakko Suominen, University of Turku, FINLAND

Technology
. Chengyan Peng, VTT Technical Research Centre of Finland, FINLAND
. Siegfried Fößel, Fraunhofer Institute IIS, GERMANY

Content & Art
. Tapio Takala, Helsinki Univ. of Technology (HUT), FINLAND
. Teijo Pellinen, Univ. of Lapland, FINLAND
. Veijo Hietala, FINLAND

CONTACT

For up to date information and further details please visit: http://www.euroitv2010.org/. Send any inquiry about the conference to: info@euroitv2010.org. To receive updates about the EuroITV2010 Conference, subscribe to the Conference mailing list by subscribing on https://listmail.tut.fi/mailman/listinfo/euroitv2010 or sending an email with "subscribe" in the subject to: euroitv2010-request@listmail.tut.fi.

Host: EMMi Lab., Department of Business Information Management and Logistics / TUT Tampere (http://www.tut.fi/emmi)

KEYNOTES

Real-time Context and Immersion Providing Uniquely new Mobile Experiences

Jyri Huopaniemi
Nokia Research Center
Tampere
Visiokatu 1
FI-33720 Tampere, Finland
jyri.huopaniemi@nokia.com

ABSTRACT
EuroITV 2010 keynote speech entitled Real-Time Context and Immersion Providing Uniquely new Mobile Experiences.

1. INTRODUCTION

Real-time context awareness has brought an entirely new dimension into mobile applications and services. The convergence of internet and mobility today enables real-time contextual information to be created and shared at any time and at any place. This real-time situational awareness combined with emerging immersive technologies such as mixed reality and 3-D rendering will provide new interactive experiences for mobile life.

In this talk, interactive solutions at the intersection of real-time context awareness and immersion will be explored. Examples of ongoing research projects will be given, and a roadmap for future research challenges will be presented.

2. SHORT BIO

Jyri Huopaniemi is Director of the Nokia Research Center Tampere laboratory. He joined Nokia in 1998 and has held various research leadership positions in Nokia in the fields of multimedia technology, mobile applications and interaction.

In his current role, Dr. Huopaniemi is leading NRC's research laboratory in Tampere, Finland that carries out strategic and long-term research on innovative media and communications solutions, including user experience and design, immersive communication, mixed reality, multimodal sensing and context, audiovisual content representation and end-to-end multimedia platforms.

Dr. Huopaniemi received his doctoral degree from Helsinki University of Technology in 1999. He is also an adjunct professor of media technology and software at the Aalto University's department of media technology.

Future Challenges for Broadcasters and the Media Industry

Marcos Gonzalez-Flower
Siemens IT Solutions and Services Ltd
68 Wood Lane
London
W12 7Tcode

Marcos.Gonzalez-Flowe@Siemens.com

ABSTRACT

Ubiquitous, cheap broadband and changing consumer behaviors are dramatically changing the way broadcasters are looking at the media industry. As Siemens travels the globe discussing the challenges its customers face, it is clear that at the forefront of there thoughts and concerns is the ultimate shape the media industry will take in the future. Using Siemens experience and the input from its customers and prospects Siemens will present a view of the future of media and the changes that broadcasters and content creators must consider if they are to survive.

Keywords

1. INTRODUCTION

The battle for the lounge is over and nobody won! The reality is that the battle for the lounge has shifted to a more appropriate battle for attention. Not only is the focus increasingly on holding a content consumers attention it is also about a delivery or distribution format that suits the consumer. The broadcasters of today are facing the greatest change to their business that they have ever experienced and instead of it coming from a change in technology such as the shift to colour from black & white, or the move to HD and 3D it is coming from a change in consumption patterns and the empowerment of consumers.

Today the biggest revenue for broadcasters continues to be advertising and linear broadcasting, with the money still rolling in it is difficult to believe that something could change this, especially for the old guard who cannot see a alternative viable business model. Having said this, the one thing that all media companies that Siemens talks with agree on is that change is coming. What the change will be and how the media industry will look are as yet uncertain, what is certain is that it won't be the same and those who don't react quickly will disappear. During this presentation Marcos will present a view of the future shape of the media industry based on the interactions he has had with media companies across the globe.

2. SHORT BIO

Marcos Gonzalez-Flower is the Global Head of Media Consulting at Siemens. Marcos runs a highly successful team of over 50 international media experts supporting media organisations around the globe ranging from the BBC in the UK to the SABC in South Africa and Corus Entertainment in Canada. With 25 years experience in the IT industry designing and implementing leading edge solutions Marcos has witnessed dramatic changes in the technologies and approaches employed to deliver services and solutions across all industries. Marcos sees the Media industry as the last great industry to truly embrace digital commodity based solutions. Driven by changes in consumer behaviour and the introduction of disruptive technologies Marcos believes the Media industry is quite possibly the most exciting place to be.

PANELS

A Panel on How the Emergence of Internet TV is Impacting the User Experience of Traditional Pay TV

Michael J. Darnell
Microsoft TV 1065 La Avenida Street
Mountain View, California, 94043,
USA +1 650 693 4965
mdarnell@microsoft.com

ABSTRACT

With the rapid growth of Internet TV (a.k.a. over-the-top TV, online TV, broadband video, etc.), many people are becoming familiar with new user experiences of video and TV consumption. Which aspects of the Internet TV user experiences are making their way into traditional pay TV? Our panel of experts will discuss and give their insights on this topic.

Keywords

Over-the-top TV, Internet TV, Online TV, Broadband video, Online video.

1. PANEL TOPIC

More and more people are watching TV content streamed or downloaded from the Internet to their computers. The user experience of watching TV on a computer differs in many aspects from watching traditional TV. Some of these differences may be perceived by people as better TV experiences. As people become familiar with Internet TV, will their expectations around user experience for traditional TV change?

One principle difference between Internet TV and traditional TV is that most of the TV programs on Internet TV are "on-demand" or watch-anytime, whereas most TV programs on traditional TV are broadcast on a schedule. Our panelists believe that the on-demand TV experience will continue to find its way into traditional TV. As Olof Schybergson, one of the panelists, said, "There is no doubt that today's child will find the concept of scheduled programming bizarre."

Another aspect of Internet TV is that it can be viewed on a variety of devices. "Consumers are more and more used to getting information on demand on any device, be it a mobile phone, laptop, their home PC or their TV set at home", said Kathrin Damian, one of the panelists. As people watch more television programs on these devices, they may expect more consistency between their Internet TV experiences and their traditional TV experiences.

Other differences between Internet TV and traditional TV experiences include the advertising models used and the richness of TV program information and enhancements on such sites as the BBC iPlayer and Hulu.

In addition, Internet TV experiences tends to have many more interactive elements than traditional TV experiences, even in the advertising. This may result in traditional TV becoming more interactive, if not directly, then on a companion device.

Our panel of experts will address these and other topics as well as questions from the audience.

PANELISTS

Jussi-Pekka Koskiranta

Cross Media Producer

YLE

Jussi-Pekka Koskiranta has long years experience in public service broadcasting and is expert in cross media productions for television.

Roger Lay

Business Consultant

SapientNitro

Roger is a business consultant specialized in digital content management and he holds a degree in general management and economics from the University of Applied Sciences Northwestern Switzerland.

Kathrin Damian

Development Manager, Digital Media Services

Alcatel Lucent

Kathrin has over 10 years in the field of interactive TV. Her previous positions include Endemic, Bertelsmann Broadband Group, RILL, Service plan Group.

Olof Schybergson

CEO

FJORD

Olof has over 10 years of experience of collaborating with major clients across the US and Europe to develop innovative service solutions and user experiences across various digital platforms.

He holds an MA in Information and Graphic Design from De Montfort University.

A Panel on the possible role of iTV in a multi-channel communication of Public Administration

Agnes Jenei Ph.D.
Corvinus University of Budapest
Faculty of public Administration
str. Ménesi 5. Budapest, 1118
Hungary +36 1 482 6654
agnes.jenei@uni-corvinus.hu

ABSTRACT

Has t-government (television-government) a realistic future outside the Big Brother of George Orwell? What kind of role the iTV can have in the multi-channel Public Administration communication? What are the characteristics of t-government in relation to that of e-government or m-government? Experts from different countries will present cases and experiences and will discuss on these issues.

Keywords

T-government, Multi-channel communication of Public Administration

1. PANEL TOPIC

Television may serve as a key medium to reach disadvantaged groups in a population. In the age of ICT, when all public service information tend to be available on-line, may iTV help delivering up-to-date public information and services into the homes of those who live on the other side of the digital divide?

The panel focuses on existing solutions and potentials of content development regarding iTV in public administration and public services. English, Brazilian, Finnish and Hungarian experts will present cases and their own experiences coming from projects and good solutions in the field of public administration and public utility services.

One of the most important aims of the panel is to demonstrate that is it possible to find a good practice of t-government approach. Guy Giles, operation manager of Looking Local platform (owned and managed by Kirklees Council, GB) will present one of the best examples in this field. For eight years Looking Local has become successful and well used government DiTV service. Nearly a hundred organizations now publish services from the realm of local governments, health, housing, police and fire authorities, creating a national, useful, easily accessible, and locally relevant portal on DiTV and on mobile phones as well.

Other experts in the panel will present more cases in the field of t-government initiatives since several projects in the world have been launched. They deploy on various interactive TV platforms along with local government organisations.

Despite the fact that pilot projects usually include highly innovating contents and citizens like them, most of the initiatives are not sustainable. Another important aim of the panel is highlighting the nature of challenges (regulatory, technology, business, providers' expertise, consumers' attitudes etc.) the realization of t-government activities face.

The panel discussion may be the first step to set an annual forum on T-Government issues in the EuroiTV conferences.

2. PANELISTS

	Guy Giles with 15 years of management consultancy, new media and software development experience in the public and private sector, is the Operations Manager for Looking Local. Looking Local is a portal on digital interactive TV, mobile phones, kiosks and games consoles, allowing access to local government information and services.
	Dr. Ari-Veikko Anttiroiko is an Adjunct Professor in the Department of Regional Studies, University of Tampere, Finland. Anttiroiko holds a PhD (Administrative Sciences) and MPhil (Philosophy) degrees, both from the University of Tampere. He is a co-editor of the "Encyclopedia of Digital Government" published in three volumes (IGR 2007).
	**Istvan Tozsa** PhD professor on e-government, heading the Public Management and Urban Studies Department and the Regional Economics Department at the Corvinus University of Budapest, having organized the FITCOM in 2004 – The First International Conference on Mobile-Government.
	Dr.-Ing. Vicente F. de Lucena Jr - UFAM/CETELI – Brazil
	**Agnes Jenei** holds a PhD in Communication Sciences. Assistant professor on communications and media studies at the Corvinus University of Budapest, having research experience and publications in different television models, innovative contents and services of interactive television.

DOCTORAL CONSORTIUM

Understanding DTV Usages in the Era of Networked Communication:

Integrating the audience, the media and the public policy's perspectives

Vera Araújo
OberCom
Palácio Foz,
Pr. dos Restauradores, Lisboa
00351 21 322 13 19
vera.araujo@obercom.pt

ABSTRACT

The transition from analogue to digital television represents a deep change, transforming television into a privileged channel to the digital world. Having in mind this role played by television, this research project departs from the following broad problematic: "How do industry, audience and policy-related features contribute to the shaping of the DTV experience?". Our main objective is therefore to analyze the role of these three dimensions in the development of DTV practices, highlighting also the interactions between these different aspects of the same reality. From a theoretical point of view, we are departing from the networked communication perspective [5], which highlights the passage from a world of mass communication (*Mass Television*), into a world of network mediated communication, combining interpersonal mediation devices with mass communication ones (*Networked Television*). On the empirical level, our research will combine both quantitative and qualitative methodologies, encompassing: 1) a survey questionnaire to the Portuguese population developed in collaboration with the Portuguese media Observatory (to research the audience side); 2) media stakeholder's interviews (to analyze the media system side); and 3) content analysis (of public policies' documents).

Keywords

Digital Television, Media Usage, Networked Communication, Domestication, Value Network.

1. OUTLOOK OF THE RESEARCH PROJECT

During the last decade, we have witnessed a vast change in the media landscape. A change, not only due to technological innovation in mediation devices themselves, but also in the ways users have chosen to socially domesticate them [5]. In this ambit, the process of television digitalization must be addressed with a special attention, as it is the medium with the widest implementation in households and as it is also a medium that is fully domesticated by individuals [18].

RESESARCH PROBLEM AND OBJECTIVES

Having in mind this role played by television, this research project departs from the following broad problematic: "How do industry, audience and policy-related features contribute to the shaping of the DTV experience?". Our main objective is therefore to analyze the role of these three dimensions in the development of DTV practices, highlighting also the interactions between these different aspects of the same reality. This broad objective unfolds itself in the following specific objectives:

- To contribute to the development of an integrated and multidimensional approach to DTV usage;
- To contribute to the understanding of DTV practices;
- To innovate discourses related to DTV, and to rethink the concept of television itself;
- To contribute for a better understanding of the challenges being faced regarding the switchover process both by the audience and by the industry;
- To contribute to a more inclusive digital TV.

MODEL OF ANALYSIS

Having in mind the case of digital television (DTV) in Portugal, the present research project aims at understanding digital media use, by combining both the perspective of the audience, and of the media system, according to a media usage model proposed by McQuail [11]. In addition, as the television switchover process is strongly marked by a specific political set of objectives, a third dimension of analysis will be added to McQuail's model: the analysis of DTV public policies.

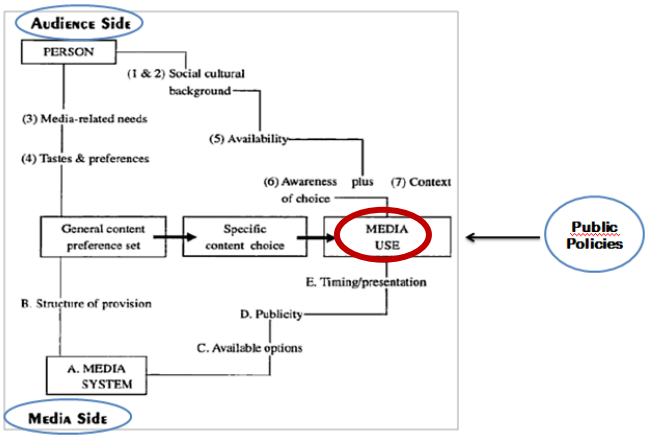


Figure. Model of Analysis (based on McQuail 2003)

THEORETICAL APPROACH

From a theoretical point of view, we are departing from the networked communication perspective [5], which highlights the passage from a world of mass communication organized by mass mediated content distribution organizations (*Mass Television*), into a world of network mediated communication still built by big media conglomerates, but also by the way people domesticate and articulate different mediated technologies, combining interpersonal mediation devices with mass communication ones (*Networked Television*). Therefore, by encompassing the audience, the media system and the public policy's perspectives, we expect to contribute to the innovation of the concept of television in the context of the global network society, departing from the assumption that the current media ecology facilitates a model away from mass communication towards a networked form of personalized mediation, and away from convergence perspectives towards media divergence approaches [5].

METHODOLOGY

On the empirical level, our research will combine both quantitative and qualitative methodologies, encompassing: 1) a survey questionnaire to the Portuguese population developed in collaboration with the Portuguese media Observatory (to research the audience side); 2) media stakeholder's interviews (to analyze the media system side); and 3) content analysis (of public policies' documents).

The next sections will provide a clear definition of objectives for each of our three research dimensions.

2. RESEARCHING THE AUDIENCE SIDE

Research on the adoption of technological innovations was initially developed in the 1980 and explores how technologies become a regular part of people's everyday lives [18]. Domestication theory in particular describes the processes by which innovations are appropriated by its users. The domestication approach considers both the practical and the symbolic aspects of the adoption and use of technologies, showing how both the meanings of things, and their materiality, are equally important in the understanding of how technologies become part of everyday life [19]. It is first and foremost a social theory as it highlights the negotiations, challenges to power and control, rule-making and breaking that accompany the introduction of technologies into any social setting.

The process of DTV domestication presents some particularities. With the introduction of DTV, rather than just being taken for granted, television has again become the centre of attention in regards to the ways we integrated it in our daily life routines, giving raise to new reflections on the role of television in the home. As stated by Gitelman [9], media systems work in two levels. The first deals with the technological specificity of the medium itself, the second, with the set of social practices traditionally associated to a given technological device. With migration to digital, television is currently suffering from an identity crisis which finds its roots precisely in the changes that are occurring at both these levels. First of all, because television isn't exclusively associated with a specific device (the television set) anymore, as audiovisual content can be broadcast to a variety of platforms. Secondly, because the cluster of social practices that used to lead the television experience is been altered. Television is a cultural form [21], a familiar object of the household, deeply entwined with the social and material practices of everyday life. Instead of projecting a completely new set of needs and wants (such as mobility and ubiquity in GSM communications), the introduction of DTV will force the culturally rooted uses of television to be greatly reinvented.

But turning an old medium to a digital, 'interactive' one seems to be even more difficult than introducing completely new technologies. In fact, as most studies on DTV domestication have highlighted, the migration to digital has brought little modifications to the television experience itself [14][20], and people tend to have a negative attitude towards DTV [10], as some have been forced to migrate due to an external push. As the implementation of DTV and DTT is first and foremost a political decision, it is somewhat perceived as an intrusive technology, failing in seducing viewers. Regarding the Portuguese literature on the subject, one should start by mentioning the reports by the Portuguese Media Observatory, Obercom, which during 2008 published a set of four brief reports about digital TV, based on the results of a quantitative survey, applied on a representative sample of the Portuguese population. Firstly, one must highlight the low percentage of people who claimed to know what the switchover is: 3,2% of the total sample. On the other hand, only 16,2% of the participants claimed that they had already heard of digital terrestrial television, which demonstrates a clear lack of audience involvement in the implementation of DTV.

Having in mind this short contextualization, we have identified the following specific objectives related to our research of the audience side:

- To perform an extensive state-of-the-art of DTV domestication research in Europe;
- To analyze the role of socio-cultural background in DTV usage;
- To examine other media-related needs and patterns of usage' impact on DTV practices, namely focusing on multitasking and networking of media devices;
- To assess the part played by personal tastes and preferences in the DTV usages;
- To understand the impact of the availability and awareness of choices in DTV practices (in terms of services offered, prices, applications, features etc.)
- To analyze the importance of the context of usage for the domestication of DTV, focusing on new contexts of reception.

As stated previously, on the empirical level our research of the audience side will be based on the development of an extensive survey designed in collaboration with the Portuguese media observatory, OberCom. Our sample will respect age, gender and region quotas. Final sample size is yet to be determined.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
EuroITV2010, June 9th-11th 2010, Tampere, FINLAND.
Copyright 2010 ACM 978-1-60558-831-5 /10/06…$10.00.

Fieldwork will be handled by a specialized market research company.

3. RESEARCHING THE INDUSTRY SIDE

The transition from analogue to digital television represents a deep change, transforming television into a privileged channel to the digital world. Throughout history, the television industry had to adapt itself to other changes [15][4]. For example, with the introduction of VHS technology, viewers were given for the first time the opportunity to control what to watch and when, giving them also the possibility to escape advertising breaks. Also, the appearance of cable television reconfigured the market, providing a much wider choice of channels and the consequent fragmentation of audiences [6][13][16].

These new consumption patterns draw new market strategies, as well as innovative market offers which will eventually lead to the development of other new consumption patterns. The traditional value chain model [17] has long been applied to the context of the study of the television sector [22][7][2]. Value creation in the ambit of television has rested, for the past decades, in a sequence of four main activities (production, aggregation, distribution and exhibition), where value was added along the chain by the different players, and where advertisers were the main founders of the system. However, this traditional concept of value is as out of time as the sequential line that characterizes value chains. Global competition, dematerialization and digitalization are in the origin of a new concept of value [3][6][4][16]. More and more, it is necessary not only to create value, but to *reinvent it* [12]. The focus of the strategic analysis is no longer the company, or the sector, but the value creation system itself, in which the different players – suppliers, consumers, business partners etc. – work together to co-produce it. In result, value chains will become progressively value *networks*, where companies and consumers concentrate not only in adding, but also in creating and reinventing value in the ambit of the Network Society [12][8][1].

Having this short contextualization in mind, our specific objectives related to the research of the media system side are:

- To assess the importance of the media system's characteristics in developing DTV practices and in creating value for the audience and the media companies;
- To analyze the role of the structure of provision in DTV practices, namely the different DTV technologies chosen to deliver television by companies (IPTV, DTT, Digital Cable);
- To understand how the publicity of the available options can influence DTV usage;
- To examine the impact of timing and presentation in DTV domestication.

As stated previously, on the empirical level our research of the industry side will be based on a set of stakeholders' interviews. The number of interviews is still to be decided. Our interviewees will be recruited among the main media and television groups in Portugal, namely: Media Capital, Impresa, RTP (all three FTA television companies), ZON (main cable company) and Portugal Telecom (main IPTV and DTT operator).

4. RESEARCHING THE POLICY SIDE

Contrary to other innovations such as for example mobile phones or iPods, the introduction of digital television is not just a market but also a political process. As the radio electric spectrum is a valuable good, its use must be maximized, and the passage of analogue terrestrial television to digital terrestrial television will allow the release of a considerable part of the spectrum, which will be made available for other applications such as mobile broadband or other TV-related features (mobile TV, HDTV etc.). Therefore, EU countries have agreed to switch-off the analogue signal in 2012, in the ambit of a broader e-society migration strategy.

In this context, several legal instruments have been developed in order to frame the TV digital switchover process, such as the alterations made in 2005 to the Television without Frontiers Directive, the Audiovisual Media Services Directive, media competition regulation and legal instruments related to the Digital Dividend. These elements are also complemented by national public policies, such as the review of the Portuguese Television Law and the launching of public consultations related to the switchover process and the digital dividend, promoted by the regulator ANACOM.

Having in mind this short contextualization, our specific objectives related to the research of the public policies' side are:

- To compare DTV related policies in EU countries and in Portugal, highlighting eventual differences and similarities, and trying to establish a typology of DTV implementation strategies;
- To analyze the priorities of DTV related policies, specifically those related to the switchover process and to the digital dividend;
- To understand the impact of those priorities in the industry and the audience in the shaping of the DTV experience.

As stated previously, on the empirical level our research of the policy side will be based on a content analysis of public policy documents related to the implementation of DTV and the switchover process.

5. REFERENCES

[1] Allee, V. (2003) *The Future of Knowledge: Increasing Prosperity through Value Networks*, Butterworth-Heinemann

[2] Andersen (2002), *Outlook for the development of technologies and markets for the European audiovisual sector up to 2010*, available at http://ec.europa.eu/avpolicy/docs/library/studies/finalised/tvoutlook/finalreport_long.pdf

[3] Bustamante, E. (2001), *La Televisión Económica*, Gedisa Editorial, Madrid

[4] Cambini, C., Valletti, T. (2002) *I mercati della comunicazione nell'era digitale*, Ed. Il Molino, Itália

[5] Cardoso, G. (2008), *From Mass to Networked Communication: Communicational models and the Informational Society*, International Journal of Communication, Vol 2 (2008), available at http://ijoc.org/ojs/index.php/ijoc/article/view/19

[6] Colombo, Fausto e Nicoletta Vittadini (eds.) (2006), *Digitizing TV – Theoretical Issues and Comparative Studies Across Europe*, Milano, Vita & Pensiero

[7] European Investment Bank (2001), *The European Audiovisual Industry*, EIB sector papers

[8] Fjeldstad, Ø., Stabell, Charles B (1998) "Configuring value for competitive advantage: On chains, shops, and networks", in Strategic Management Journal nº19

[9] Gitelman, L. (2006), *Always Already New: Media, History, and the Data of Culture*, MIT Press.

[10] Impulsa TDT (2008), *Anuário TDT 2007*, available at http://www.impulsatdt.es/pdf/ponencias/ANUARIO_TDT_2007.pdf

[11] McQuail (2003), *Teoria da Comunicação de Massas*, Gulbenkian, Lisboa.

[12] Normann, R. and Ramírez, R. (1993) *From Value Chain to Value Constellation: Designing Interactive Strategy*, Harvard Business Review, 71, July/August, 1993, pp. 65-77

[13] Obercom (2007), *Contributos para uma análise da implementação da TDT*, disponível em http://www.obercom.pt/client/?newsId=373&fileName=fr3.pdf

[14] Ofcom (2008), *Switchover Progress Report*, available at http://www.digitaluk.co.uk/__data/assets/pdf_file/0011/6500/trQ407.pdf

[15] Papathanassopoulos, S. (2002), *European television in the digital age*, Polity Press, UK

[16] Picard, R. and Brown, A. (eds) (2004), *Digital Terrestrial Television in Europe*, USA.

[17] Porter, Michael E., "Competitive Advantage". 1985, Ch. 1, pp 11-15. The Free Press. New York.

[18] Silverstone, R., Haddon, L. (1996), "Design and the domestication of information and communication technologies: technical change and everyday life", in Mansell, R., Silverstone, R (Eds),*Communication by Design: The Politics of Information and Communication Technologies*, Oxford University Press, New York, NY, pp.44-74.

[19] Silverstone, Hirsch and Morley, (1992) 'The Moral Economy of the Household' in, Silverstone R., & Hirsh, E. (1992). Consuming Technologies: Media andInformation in Domestic spaces London: Routledge.

[20] Van Dijk,J. and A. Heuvelman, O. Peters (2003) 'Interactive television or enhanced television? The Dutch users interest in applications of ITV via set-top boxes' Paper presented at the 2003 annual conference of the International Communication Association in San Diego USA

[21] Williams, R. (1974), *Television: Technology and Cultural Form*, Routledge.

[22] Zerdick A. et alia (2000), *E-conomics : strategies for the Digital Marketplace*, Springer, Berlin

Converged Mobile Media: Evaluation of an Interactive User Experience

Alexandre Fleury
Aalborg University
Dpt. of Electronic Systems
Niels Jernes vej 12
9220 Aalborg, Denmark
amf@es.aau.dk

ABSTRACT
This paper describes a PhD thesis exploring various aspects of the end user experience with mobile rich media services. First, the author briefly introduces the background fields that frame the study. Three research questions are then formulated and their scientific contribution is justified. Subsequently the author proposes a methodology to investigate these questions and reports on two initial user studies. Finally, the thesis' ongoing work and intended future steps are introduced.

Categories and Subject Descriptors
H.5.2 [**Information interfaces and presentation (e.g., HCI)**]: User Interfaces – *evaluation/methodology*

General Terms
Human-computer interaction, convergence, interactivity

Keywords
User studies, mobile television, study plan

1. INTRODUCTION
This thesis explores the potential of mobile devices to support the shift from fixed to mobile in media related practices within today's converged media landscape. The thesis initially argues that this shift can be successfully achieved if the concepts of mobility and interactivity are naturally introduced in the context of media consumption. The thesis will therefore investigate how the establishment of the mobile device as a complement to home media equipment impacts the perception and usage of media services. In particular the research pursued throughout this project will examine the development of the Danish mobile television and associated services as its main case study. In fact, the thesis takes place in the context of the Converged Advance Mobile Media Platforms (CAMMP) project[1], which addresses the convergence of media services with mobile technologies. In this purpose, the project partners merge mobile technologies with Internet, digital TV and radio and investigate the potential of the resulting infrastructure that combines traditional media and user-generated content.

1.1 Previous Research on Mobile TV
A large body of research has been dedicated to the various aspects of mobile media in general and mobile television in particular since the early 2000s. However, only few studies involving users have been conducted in Denmark so far, due to the lack of infrastructure enabling such tests. This thesis will try to fill in this gap by not only verifying previous results and their applicability to the Danish environment, but also investigate uncovered issues within the fields described in details in section 2.

One of the first large scale studies about mobile TV involving users has been conducted in 2003 as part of the Finnish Mobile TV project [15]. This comprehensive study covered topics from the possible content of mobile TV to design requirements for the graphical user interface (GUI). The participants in the study responded in general positively to the concept of watching mobile television on a mobile device, however they expressed their concern about the usability of watching video content on a small screen. This concern has been extensively investigated by H. Knoche and his colleagues. For instance in [10], they investigated image resolution and bitrate requirements for delivering an acceptable watching experience on 3.5 inches displays; and in [11] they focused on the effect of zooming on the perceived quality of experience (QOE).

While early studies focused on the challenges of watching television content on mobile devices, more recent studies have been increasingly focusing on the mobile and social aspects of watching TV on the move. As explained in section 2, this PhD thesis follows these tracks.

2. FIELDS OF STUDY
Three general themes ground the dissertation's research questions based on an initial literature review. This grounding material provides opportunities for not only validating and extending previous results, but also for studying new issues via a cross-disciplinary approach. The literature study focuses on the themes introduced in the following sections.

2.1 Mobility
Within this theme, focus is put on the study of the contexts in which media content is consumed on mobile devices. The objective is to build on previous work by for instance J. Chipchase and his colleagues [4], who investigated four primary contexts in which mobile TV is consumed. The PhD study will mainly focus on the home, macro-break and public transportation scenarios in the Danish context.

Inferring contexts from data collected on the participants' phone during longitudinal studies is also a critical aspect of such research. Many tools have been designed for recording and processing contextual data but there is a noticeable lack of standardization and common agreement on methods and tools to be used in such studies [6]. Input in that domain will be provided via collaboration with senior researchers in this area.

[1] http://www.cammp.aau.dk/, April 2010

2.2 Interactivity

Focus within this theme is put on the interaction between end users and mobile media related products (devices or services). The starting point for this study is the previous research on the second screen approach reported in [2], which investigated peripheral activities TV viewers engaged in while consuming television on their home equipment. Findings from this experiment and from other similar ones will be further developed, mainly through scenarios involving content transfer between fixed and mobile equipment, as well as manipulating the electronic program guide (EPG) while watching a TV program.

2.3 Usability Evaluations

Mobility arguably introduces constraints on the design and execution of user studies. This part of the review builds on the work reported in [9], targeting methods and tools dedicated to field and lab trials and their suitability for mobile media studies. In particular the interactions between researchers and test participants and the involvement of users at various stages of the experiment are to be scrutinized. Since the whole PhD project relies on conducting such user studies, it is intended to follow the theories and guidelines introduced in [12] and [15].

3. RESEARCH QUESTIONS AND METHODOLOGY

Three research questions have been identified as of primary interest for the thesis. As detailed in this section, the contributions to the research community evolve with the questions from general to specific.

1. How does mobility impact media consumption and creation in a society that is always connected

Answering this question will contribute to the validation of new HCI patterns in mobile media consumption. The expected outcome is a deeper understanding of the effects of the ubiquitous access to digital media on consumption and creation practices.

In order to answer this question, a hierarchy of situations in which mobile media is consumed will be created. Such situations can be characterized by a combination of dimensions, including the user's location and activity, the surrounding environment, etc. Then it is intended to study the evolution of media usage through in-depth interviews and remote logging of media consumption during field trials. Finally, a comparative study of common practices in two different markets (namely Denmark and Japan) will be performed through in-depth interviews with senior researchers, as well as active participation in research projects in both markets.

2. How do end users perceive and use the combination of mobile and fixed devices in their media experience?

This question aims at informing how consumers interact with multiple pieces of equipment when dealing with media. Emphasis is put on the usability of transferring content and manipulating services in a multi-devices environment.

This question will be tackled through the investigation of (1) the handover experience while transferring content between devices or varying nature, (2) the mobile phone as a second screen for interacting with television programs and services, (3) the mobile phone as a universal remote enabling the control of various home media appliances. Furthermore, it is intended to assess the acceptability of seamless integration across services through contextual data collection from field trials and in-depth interviews.

3. How to improve the researcher/test user communication and measure the test user's reliability?

This last research question addresses specific issues related to evaluation methods involving users. It is intended to answer the question of reliability by comparing media consumption intentions as expressed by users and their actual consumption when using a fully implemented service in their everyday life. Using multimodal analysis techniques during usability evaluations will also help assessing the reliability of test participants. Additionally, the impact of recurrent (educated) test participants on the evaluation results will be measured.

For what concerns the communication between researchers and test participants, the focus will be put on large scale unsupervised studies. Various communication schemes will be evaluated and compared. Finally, a cost benefit analysis will be conducted by comparing the results collected from various test settings and linking them to the cost of implementing the test setup.

4. PRELIMINARY RESULTS

Two user tests have been conducted so far in the context of CAMMP and documented in [4]. The studies aimed at defining basic knowledge about how Danes relate to mobile TV and provide indication about which specific areas should be further investigated in priority. Focus was put on (1) the acceptability of consuming mobile television in a social environment and (2) users' collaborative and competitive behaviors with regards to mobile media.

4.1 Situated Task-Based Interviews

The purpose of this first test activity was to investigate how users handle the consumption of rich media in a social context, covering the following questions:

- Is it natural for users to consume rich media in a social context?
- Do issues such as privacy and personal sphere play a role in such contexts?
- Does the social context with its inevitable auditory/visual impact call for specific needs, such as head-phones to cancel ambient noise?
- How does channel switching perform on the mobile device?

This first study showed that users would feel comfortable watching TV on their mobile phone when surrounded by strangers. However the use of earplugs would be necessary to ensure privacy and an appropriate experience.

4.2 Group Discussions

The purpose of this second activity was to investigate the concepts of competition and collaboration involved when using rich media services in a social mobile context. The results from this evaluation activity provide a basis for future elaboration on interactive rich media scenarios, associated services and their context of usage. Additionally, the study uncovered issues to be further investigated in order to ensure mobile participation to a collaborative service. Such issues include the rewards and the information quality.

5. CONCLUSIONS

This paper introduced a doctoral dissertation investigating various aspects of user interaction with mobile rich media systems. Three research questions relevant to the field of interactive television have been formulated and a list of intended user tests has been established. The following paragraphs conclude the presentation of the thesis by firstly summarizing the thesis' scientific contribution and secondly describing the future work being currently planned.

5.1 Contribution Summary

The three areas covered by this thesis (mobility, interactivity and usability evaluation), by answering the three research questions stated in section 2, will provide the scientific community with insights on key aspects of the user experience with mobile media service, and especially mobile interactive television. The challenges inherent to mobile services such as the editorial format of the content, the limited screen size or the effects of network coverage will be informed by end users. The enhancement of the fixed television experience by the introduction of services and devices allowing cross-platform scenarios will be scrutinized through the design and evaluation of prototypes. Finally, the involvement of participants in user studies and their influence on the results will be investigated and documented in order to facilitate further research involving test participants.

5.2 Current and Future Work

Currently the author is involved in designing the following test activities as part of CAMMP:

- Seamless handover experience across devices. It is intended to investigate how people perceive the transfer of content from a mobile phone to a television and the preferred set of actions to perform the transfer.
- Determination of acceptable content switching delays. The purpose of this activity is to assess the acceptability of various delays while switching content on a mobile device under various contextual conditions.
- Confirmation of specific content needs. This test aims at validating the results collected in previous tests with regard to the type of content wished by end users.

In addition to planning these test activities, the author is currently engaged in an extensive literature review on the themes of interactive television and especially the second screen approach. Concerning the former it is intended to review the work by J. F. Jensen (e.g. [6], [8]), which covers general aspects of interactive TV, while the latter concerns primarily the work by P. Cesar (e.g. [1], [3]) and A. C. Roibás (e.g. [13], [14]). Additionally, the work published in the proceedings of past EuroITV conferences is reviewed on a continuous basis.

Finally, the author has planned interviews and is conducting shared projects with colleagues from other departments at Aalborg University. Additionally, a visit to a Japanese research institution is to be confirmed in order to not only broaden the thesis' perspective but also to investigate specific issues of common interest in the field of mobile media creation.

6. ACKNOWLEDGMENTS

The author would like to thank the CAMMP project for supporting the work conducted as part of this thesis.

7. REFERENCES

[1] Cesar, P., Bulterman, D. C., Jansen, J., Geerts, D., Knoche, H., and Seager, W. 2009. Fragment, tag, enrich, and send: Enhancing social sharing of video. ACM Trans. Multimedia Comput. Commun. Appl. 5, 3 (Aug. 2009), 1-27. DOI= http://doi.acm.org/10.1145/1556134.1556136

[2] Cesar, P., Bulterman, D. C., and Jansen, A. J. 2008. Usages of the Secondary Screen in an Interactive Television Environment: Control, Enrich, Share, and Transfer Television Content. In Proceedings of the 6th European Conference on Changing Television Environments (Salzburg, Austria, July 03 - 04, 2008). M. Tscheligi, M. Obrist, and A. Lugmayr, Eds. Lecture Notes In Computer Science, vol. 5066. Springer-Verlag, Berlin, Heidelberg, 168-177. DOI= http://dx.doi.org/10.1007/978-3-540-69478-6_22

[3] Cesar, P. and Chorianopoulos, K. 2009. The Evolution of TV Systems, Content, and Users Toward Interactivity. Found. Trends Hum.-Comput. Interact. 2, 4 (Apr. 2009), 373-95. DOI= http://dx.doi.org/10.1561/1100000008

[4] Cui, Y., Chipchase, J., and Jung, Y. 2007. Personal TV: a qualitative study of mobile TV users. In Proceedings of the 5th European Conference on interactive Tv: A Shared Experience (Amsterdam, The Netherlands, May 24 - 25, 2007). P. Cesar, K. Chorianopoulos, and J. F. Jensen, Eds. Lecture Notes In Computer Science. Springer-Verlag, Berlin, Heidelberg, 195-204.

[5] Fleury, A., Pedersen, J. S., and Larsen, L. B. 2009. Two perspectives on mobile television: Consumption in a social context and collaborative/competitive behaviors. In Proceedings of the Ninth Danish Human-Computer Interaction Research Symposium, (December 2009), 18-21.

[6] ter Hofte, H., Jensen, K. L., Nurmi, P., and Froehlich, J. 2009. Mobile Living Labs 09: Methods and Tools for Evaluation in the Wild: http://mll09.novay.nl. In Proceedings of the 11th international Conference on Human-Computer interaction with Mobile Devices and Services (Bonn, Germany, September 15 - 18, 2009). MobileHCI '09. ACM, New York, NY, 1-2. DOI= http://doi.acm.org/10.1145/1613858.1613981

[7] Jensen, J. F. 2005. Interactive television: new genres, new format, new content. In Proceedings of the Second Australasian Conference on interactive Entertainment (Sydney, Australia, November 23 - 25, 2005). ACM International Conference Proceeding Series, vol. 123. Creativity & Cognition Studios Press, Sydney, Australia, 89-96.

[8] Jensen, J. F. 2008. Interactive Television - A Brief Media History. In Proceedings of the 6th European Conference on Changing Television Environments (Salzburg, Austria, July 03 - 04, 2008). M. Tscheligi, M. Obrist, and A. Lugmayr, Eds. Lecture Notes In Computer Science, vol. 5066. Springer-Verlag, Berlin, Heidelberg, 1-10. DOI= http://dx.doi.org/10.1007/978-3-540-69478-6_1

[9] Kjeldskov, J., and Graham, C. 2003. A Review of Mobile HCI Research Methods. In Human-Computer Interaction with Mobile Devices and Services, vol. 2795/2003. Springer Berlin/Heidelberg, 317-335.

[10] Knoche, H., McCarthy, J. D., and Sasse, M. A. 2005. Can small be beautiful?: assessing image resolution requirements for mobile TV. In Proceedings of the 13th Annual ACM international Conference on Multimedia (Hilton, Singapore, November 06 - 11, 2005). MULTIMEDIA '05. ACM, New York, NY, 829-838. DOI= http://doi.acm.org/10.1145/1101149.1101331

[11] Knoche, H., Papaleo, M., Sasse, M. A., and Vanelli-Coralli, A. 2007. The kindest cut: enhancing the user experience of mobile tv through adequate zooming. In Proceedings of the 15th international Conference on Multimedia (Augsburg, Germany, September 25 - 29, 2007). MULTIMEDIA '07. ACM, New York, NY, 87-96. DOI= http://doi.acm.org/10.1145/1291233.1291254

[12] Nielsen, J. 1994. Usability Engineering. Morgan Kaufmann.

[13] Roibás, A. C., Geerts, D., Furtado, E., and Calvi, L. 2006. Investigating new user experience challenges in iTV: mobility & sociability. In CHI '06 Extended Abstracts on Human Factors in Computing Systems (Montréal, Québec, Canada, April 22 - 27, 2006). CHI '06. ACM, New York, NY, 1659-1662. DOI= http://doi.acm.org/10.1145/1125451.1125756

[14] Roibás, A. C. and Johnson, S. 2006. Unfolding the user experience in new scenarios of pervasive interactive TV. In CHI '06 Extended Abstracts on Human Factors in Computing Systems (Montréal, Québec, Canada, April 22 - 27, 2006). CHI '06. ACM, New York, NY, 1259-1264. DOI= http://doi.acm.org/10.1145/1125451.1125686

[15] Rubin, J. 1994 Handbook of Usability Testing: how to Plan, Design, and Conduct Effective Tests. John Wiley & Sons, Inc.

[16] Södergård C. (ed.). Mobile television - technology and user experiences, Report on the Mobile-TV Project. Espoo: VTT Publications 506, 2003.

Secondary-screen as a control device for TV content - Measuring user behavior.

Francisco Javier Burón Fernández

Associated Lecturer in Cordoba University (UCO)

Campus Universitario de Rabanales
CP 14014, Córdoba (Spain)
+34957211020

jburon@uco.es

ABSTRACT
One relevant research issue is if delegating advanced visual interface between a Tablet and a TV could compete or cooperate for the benefit of the user. Moreover, learning more about how multi-device interaction is appreciated in multi-users. Previous research has raised and considered the above research issues and questions for dual screen set-ups in the work environment. Broadly it is explored multi-device user interface configurations in the context of a leisure environment and for entertainment applications. Our objective is to provide interaction possibilities that are more than the sum of the parts.

Categories and Subject Descriptors
H.5.2 [**User Interfaces**]: Graphical user Interfaces - Input devices and strategies - Interaction styles

General Terms
Design, Human Factors.

Keywords
HCI, Interactive Television (iTV), Remote Controller.

1. INTRODUCTION
One of the less desirable effects of the introduction of information technologies and communications on the Internet and the Web is information and content overload as well as the lack of accessibility, usability and adaptivity of the interfaces of electronic devices, applications and digital content so that they comply with regulations concerning accessibility and de facto and de jure standards.

Undoubtedly, the interactive television allows us new ways to interact with the Internet creating a drastic change in human-computer paradigm. Moving various Internet services to television is a challenge for the user experience aimed at achieving customer's satisfaction in the interaction with the system regardless of disability presented (Jensen, 2005).

Dan O'Sullivan defines iTV as a oxymorron (Cesar et al., 2009) and its great potentiality is also the biggest problem for standardization, definition, conceptualization and development.

SIeSTA is a Web Operative System supported by Linux targeted handicapped people to help them to access to technologies, e-services, helping technologies and trough iTV. The SIeSTA system wants to be a step forward in the domain of interfaces. The information currently represents this domain is perhaps the most dispersed, heterogeneous and fragmented that exist on the Internet, so we need a process of integrating this content into a single system accessible and usable through a single multimodal interface that facilitate access to them.

Currently, a remote controller controls SIeSTA system. The objective of the PhD is to develop and evaluate alternatives scenarios of new interfaces based on dual screen to interact with this iTV system.

While there is research that evaluates the usability or performance of the independent displays as a single continuously addressable space, there is also research that employs two synchronized screen devices. Both areas of study are of great interest and influence in our research for both the evaluation of its use and for the development of our prototypes. In the following paragraphs, we describe previous work that has used dual displays.

2. RELATED WORKS
The majority of previous research in dual-screen set-ups has been focused on the effects of increased screen real-estate, which has been considered as a quantitative parameter in performing several user tasks. Indeed, several studies have been performed in a work-setting, which lends itself to performance measurements as efficiency. On the other hand, there are few research efforts in leisure environments that have considered the qualitative effects of secondary screens. The general research area is characterized by the partnership between a personal device and shared screen one. Then, one significant research issue is to balance the visual interface system between two devices with output abilities.

2.1 Multi-device timeline
In a parallel line with this research we have found various studies that make use of a second screen device to interact with iTV.

Since the advent of the PDAs there have been some studies to replace the remote control in the interaction with interactive television. One of the most influential research for this work is the Robertson one (1996), which proposes a prototype for real estate searching by a PDA bidirectionally communicated via infrared with interactive television. The author proposes a design guide remarking the importance of distributing information through appropriate devices. So the right information for display on PDA's

is text and some icons, but television is suitable for displaying large images, video or audio. So the nature and quantity of information determines how to display and on which device. This research also gives priority to increase a synchronized cooperation between both devices.

In the design proposed by Sanaz (2005) is established a mobile phone as main element of interaction with the television in order to language learning. This interaction, unlike the previous one, is based on multiclient-server arquitecture: mobile phones connect through two levels of WAP and SMS to the server, which is accessed via a set top box for the iTV.

Pablo Cesar et al. (2009) propose a more complex model arquitecture. Their work defines taxonomy of global actions with interactive television. This classification is divided into three levels high: content control, content enrichment and content sharing. In this paper we focus our prototypes on content control, covering among other things navigating of content, and the common video controls (Play, Pause, etc). Also we provide some prototypes for enrichment content.

2.2 Dual-screen research

In the study area of multiple-screens we should emphasize the study of Hutchings (2004) where is compared the general use of a single with multiple monitors (2 or 3). To do this, is evaluated how 31 people use different windows on Windows XP operating system. To carry out the evaluation is used a monitoring tool. As highlight features we can observe that the activation of additional screens arises as a consequence of the desire to hide information in the main screen display. Besides it is important to highlight that is not usually interact with the various windows of the same program.

Another study to really consider is from Grundin (2001). It shows that the users do not treat the second monitor as an additional space, so not establish a single window across multiple monitors. In addition users will typically set a monitor for the primary task and other tasks related to but not synchronized. Like other studies confirm the increased usability of multiple monitors in achieving greater satisfaction of users and more productivity.

3. RESEARCH QUESTIONS

In summary, the PhD is motivated by the introduction and wide adoption of small and powerful mobile computers, such as smart phones and tablets. The latter has raised the opportunity of employing them into multi-device scenarios and blending the distinction between input and output. In particular, we are addressing the following research questions: Do the advanced visual interfaces compete or cooperate for the attention and the benefit of the user? Most notably, how coupled-display visual interfaces are appreciated in multi-user scenarios?

4. METHODOLOGY

In this research, we are exploring multi-device user interface configurations in the context of a leisure environment and for entertainment applications.

4.1 Technological Set-up

The arquitecture for this prototype consist of:

- A TV connected to a set-top-box based on Linux.
- A tablet with Linux operating system installed (Figure 1)
- A local network that it is connected both devices.
- A remote controller connected to the set-top-box using Bluetooth. The design of the controller is based on a selection system based on six colors (Figure 2). These colors correspond to different options in the menu.

Figure 1 Tablet with SIeSTA system

One of the most important settings is the way in which the tablet is connected bidirectionally with the set-top-box. To connect both interfaces we could use Bluetooth, RPC or HTTP-Request. Bluetooth is already used to interact between remote control and iTV. However, it is a better option to use RPC for a complex interaction because it is more scalable and flexible. For some operations and transmission of content is used HTTP-request. So, in the end, we can define the arquitecture as a double client/server between the tablet and the set-top-box. To implement the interface has been chosen HTML5 for its opened character.

Figure 2 - Current SIeSTA remote controller

4.2 User Evaluation

Dual-screen interaction might not be suitable for every type of Television content. Actually, it might be rather suitable for some types of content, but completely irrelevant for other types of content. Although researchers have highlighted some of the benefits (e.g., personalized view of related content), they have not yet coupled them to the respective types of content. Instead, previous efforts have only regarded the technological facts, such as the segmentation of long videos, in shorter clips and providing links to related information. As a matter of fact, obvious choices of dual-screen compatible content include sports, news, documentaries, series, and movies.

This work is focused on the evaluation of a secondary-screen as a control device for TV content. Previous research has regarded the secondary-screen as an editing and a sharing interface, but has neglected the control aspect. Moreover, previous research has only concerned user attitude, but has neglected to employ a

methodology that explains actual user behavior. Indeed, Cesar et al. (2008) have focused on the utility and the general acceptance of a dual-screen system, but have not employed any user behavior measurements. In particular, we are seeking to understand the balance between the shared and the personal screen during alternative TV-control scenarios that regard the secondary-screen as a: 1) simple remote control, 2) related information display, 3) mirror of the same TV content.

The main objective in the evaluation of a dual-screen TV set-up is the measurement of actual user behavior rather than just user attitude. For this purpose, we are measuring user attention and engagement with TV content. In contrast to measurements of efficient and effective task completion, which are common in work settings, we are focusing on measurements of user involvement with the TV content, which are common in a leisure setting. Moreover, we are working on measurements that consider the main of benefit of TV, which can be summarized as "a significant shared experience" within smaller or larger social circles, and regardless of the actual or perceived quality of the content.

In particular, we have developed a flexible experimental set-up for testing several hypotheses, such as those developed in previous related research:

- Cesar et al. (2009) 'argue that secondary screens provide a less obtrusive mechanism for affecting television content than traditional solutions in the form of television overlays.'
- 'A number of participants did not want to browse while something was already showing' (Cesar et al. 2008).

4.3 Outline of ongoing research

For our research we consider the following situation: Peter is watching the TV on demand and he wants to control the video content (play/pause/stop) and do some interactive actions like: see more information about the video, mark as favorite, share and comment it and see related videos.

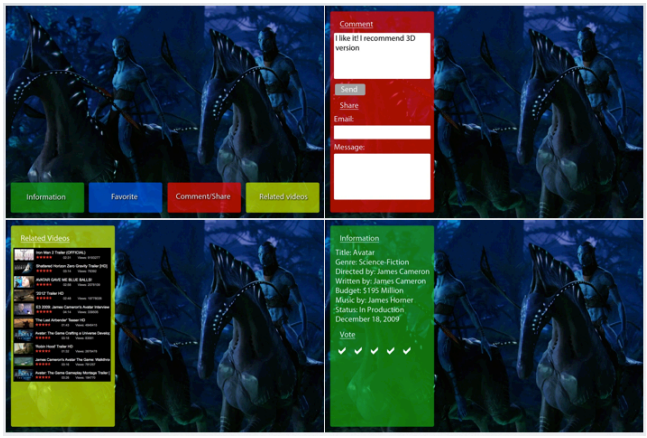


Figure 3 - Scenarios 1 and 3

For this situation we have proposed three scenarios of iTV:

1. To Interact with iTV using a remote control (Figure 3): In this case user interacts with iTV using remote controller (Figure 2). To control the content there is a button in the remote controller to play or pause the video. To use interactive actions: Information, Favorite, Content/Share and Related Videos, the user press the color buttons to access every one. When the user presses one-color button a bigger rectangle is opened. Now the remote controller is used to move (up or down) the focus into the content. When the user wants to select an option he would press the central button. To introduce text it is necessary a extra keyboard.

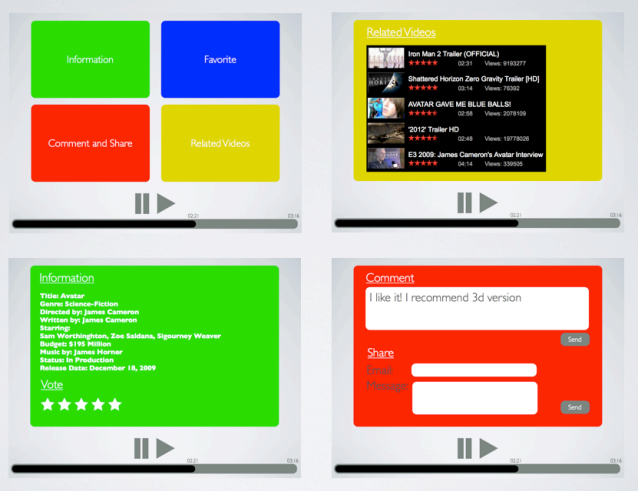


Figure 4 - Scenario 2

2. To interact with iTV using a tablet as remote controller (Figure 4): In this case, all the overlay information shown in the first scenario is displayed in the tablet cleaning the first screen of interactive information so it wouldn't disturb other users. Now to introduce text is shown a virtual keyboard in the tablet screen. (Figure 5)
3. iTV inside the tablet (Figure 3): This scenario suppose that the user is watching the iTV in the tablet so the prototypes is very similar to the Figure 3 but, in this case, it would be the tablet screen. In this prototype would be necessary introduce a video control bar similar to the scenario two (Figure 4)

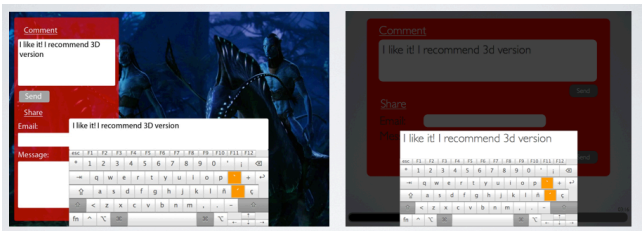


Figure 5 - Virtual Keyboards in scenarios 1 and 3 (left) and in scenario 2 (right)

5. RESEARCH PLAN

The research plan is addressing the following questions: Do the advanced visual interfaces help the user? Does the user loose his attention with these interfaces? Most relevant, how coupled-display visual interfaces are appreciated in multi-user scenarios? It is expected that the case study of TV users and TV content could provide complementary evidence for the design of coupled display interfaces in general.

6. ACKNOWLEDGEMENTS

The author wishes to acknowledge the help of Enrique García Salcines, Carlos de Castro and Konstantinos Chorianopoulos in bringing their experience.

7. REFERENCES

[1] Cesar, P., Bulterman, D. C., Geerts, D., Jansen, J., Knoche, H., and Seager, W. 2008. Enhancing social sharing of videos: fragment, annotate, enrich, and share. In Proceeding of the 16th ACM international Conference on Multimedia MM '08. ACM, New York, NY, 11-20.

[2] Cesar, P., D.C.A. Bulterman, and J. Jansen, 2009.Leveraging the User Impact: An Architecture for Secondary Screens Usage in an Interactive Television Environment. Springer/ACM Multimedia Systems Journal (MSJ), 15(3): 127-142

[3] Cesar, P. and Chorianopoulos K. (2009). The Evolution of TV Systems, Content, and Users Toward Interactivity. Foundations and Trends in Human–Computer Interaction, 2(4): 279–374.

[4] Fallahkhair, S., Pembertom, L. and Griffiths, R. 2005. Dual Device User Interface Design for Ubiquitous Language Learning: Mobile Phone and Interactive Television

[5] Grudin, J. 2001. Partitioning digital worlds: focal and peripheral awareness in multiple monitor use. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '01. ACM, New York, NY, 458-465.

[6] Hutchings, D. R., Smith, G., Meyers, B., Czerwinski, M., and Robertson, G. 2004. Display space usage and window management operation comparisons between single monitor and multiple monitor users. In Proceedings of the Working Conference on Advanced Visual interfaces. AVI '04. ACM, New York, NY, 32- 39.

[7] Jensen, J.F. (2005) Interactive Television: New Genres, New Formats, New Content. In: Austral- asian Conference on Interactive Entertainment. Sydney

[8] Kang, Y. and Stasko, J. 2008. Task/application performance using single versus multiple monitors: a comparative study. In Proceedings of Graphics interface 2008. GI, vol. 322. Canadian Information Processing Society, Toronto, Ont., Canada, 17-24

[9] Robertson, S., Wharton, C., Ashworth, C., and Franzke, M. 1996. Dual device user interface design: PDAs and interactive television. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '96. ACM, New York, NY, 79-86.

[10] Weiser, M. Some computer science issues in ubiquitous computing. Communicationsof the ACM, 36, 7, (July 1993), 75-85.

Augmented local television and visual knowledge building
Shared space learning objects and interactive services in a multiuser mixed reality
Research Plan.

John Grönvall
Arcada University of Applied Science
Vadgränd 1c31 00200 Helsinki +358 50 368 4607
john.gronvall@arcada.fi

ABSTRACT
The aim of this work is to create and assess a new system for visual knowledge building. The system core is a local television cable channel augmented with a novel type of visual knowledge discovery interfaces. These interfaces convert shared space physical objects and locations into learning objects by linking social video material and augmented reality based overlays to the object P thus creating an interactive media interface to the World Design Capital of Helsinki 2012. In the empirical part of my work I will analyse and assess the way the media interface is used and how it is perceived as a knowledge building and opinion making tool P both among "digital natives" and late adaptors. My primary research method is phenomenographical complemented by a case study. The principal relevance of this dissertation - from a pedagogic point of view, is the value of the resulting new didactical tools and the knowledge and inspiration the convey. From a media-technological perspective the system created can be utilized as a platform forming a base for sensory multiuser games and other media concepts. From a sociological view, the results will empower the potential citizen journalists of Helsinki and give them a medium to express themselves in. Finally the thesis generates proof of concept for a general citywide media infrastructure that could be replicated in other major capitals.

Supervisor: professor Stefan Sonvilla-Weiss, Aalto University Helsinki

Categories and Subject Descriptors

H.1.2 [**User/Machine Systems**] *Human factors.*

General Terms
Algorithms, Management, Documentation, Performance, Design, Experimentation, Human Factors, Theory, Legal Aspects, Verification.

Keywords
Digital services, media ecology, knowledge building, citizen journalism, knowledge transfer, transdiciplinary, context dependent, social video, gaming, barcodes, NFC, multi-camera production.

1. INTRODUCTION

Until recently the concept of television has been quite straightforward, however during the last two decades when different media truly have started to converge the word "television" is no longer a well-defined, precise notion or concept (Storsul, Stuedahl 2007). The implications of this paradigm shift, where large media houses are being challenged by individual media producers, is especially interesting in that it opens up new possibilities for a truly democratic local media environment in the form of citizen generated content and DIY media production.

1.1 Local TV – the Stadi.TV platform

Local television has been around almost as long as TV itself. Smaller cities in the Swedish speaking parts of Finland (e.g. Ekenäs and Närpes) have a long tradition of local content being produced and broadcasted via cable. In contrast, few of the larger cities in Finland have their own local-TV stations. At Arcada University of Applied Sciences in Helsinki we have, for didactical reasons, founded a nonprofit cable-TV channel called DINA. During the last four years we have been working as a kind of mash up enabler - developing new tools and services in order to facilitate creation of mash up media production services. Notably the massive project around the Eurovision song contest in 2007 has given us valuable insight into new ways of helping users produce content themselves.

The experiments and tools we have created have formed an interest among the city municipals and larger telecom and research companies. We have therefore received funding to expand on the traditional TV-channel to include a multitude of new services on the web and for the mobile phones. This next iteration of DINA is called Stadi.TV – a Local TV station in Helsinki that is endorsed and funded by the city municipals themselves. The channel will be launched on September 29, 2010 – with the objective to give the citizens of Helsinki an opportunity to get involved in content production and digital democracy.

The Stadi.TV framework is also an important factor in the research project *Social Video* - a part of the ICT cluster of the Finnish Strategic Centres for Science, Technology and Innovation (ICT SHOK) coordinated by TIVIT Oy. As a researcher and representative of Arcada in the SHOK programme, I participate in creating new TV production tools and services and expand on the ways in which content is produced and delivered for television and other media platforms – such as the web and mobile handheld devices.

The same organisations that form the Stadi.TV consortium will have a major role in developing the media surfaces of Helsinki in

preparation for the design capital of the world year of 2012. Herein lies my interest in revisiting the phenomenon of local TV and augmenting it with new media-technological receptacles that will lure the visitors of the design capital into a deeper understanding of what the city has to offer.

1.2 Conceptual framework

The conceptual frameworks that I will start out with and later hopefully develop into a full theory, is loosely based on the work of the Media Ecology founders by McLuhan (1964), Postman (1985) and Ong (1982). It is the study of media environments, the idea that technology and techniques, modes of information and codes of communication play a leading role in human affairs.

I'm also interested in utilizing the "proxemic laws" theories Edward T. Hall. (1963)

I will use some of the newer reincarnations of the notion of modernism such as Liquid modernity as defined by Bauman (2000). This is a suitable framework for discussing the now intertwined roles of the producer and consumer in a local TV environment.

I'm very interested in the notions of mediation and mediatization. Studying the effects of the substitution of the real with a mediated version will inevitably force me into the field of postmodern media theories.

The role of the citizen as producer and consumer can also be discussed via the concept of participatory journalism (Shaun and Willis 2003) – an umbrella term describing the forms of blogs, online communities and social video. Furthermore I will also discuss whether the concept of media convergence can be applied to my research (Jenkins, 2006).

Finally I will be using the term Digital culture (Gere, 2006) as a means for discussing the relationships between the arts, culture and technology.

1.3 Research goals

In my research I will combine four related areas that together form a new technological system that will alter the traditional way of doing local TV and thus create a new media platform. These areas are: enabling technologies (barcodes, phones, AR-technologies), production system design (defining metadata handling, building broadcast backend), production principles (best practices and new way of using production tools) and programme concepts (new "TV" formats and ideas).

The system core is a local television cable channel augmented with a novel type of visual knowledge discovery interfaces. These interfaces convert shared space physical objects and locations into learning objects by linking social video material and augmented reality based overlays to the object – thus creating an interactive citywide media interface.

These learning objects can be any surface, object, service or place in the city – by following a link presented via a 2D bar code or a NFC-tag, the agent is transferred onto the web where he can consume or produce video material. Furthermore – if there is time I would like to enhance the experience by also linking a Augmented Reality display on top of whatever video material the person is watching.

Once the conceptual model has been pilot-tested the next step is to engage both professional media producers and the citizen in creating informative video material that can show off Helsinki as a tourist friendly capital, with advanced local services both for digital natives and late adaptors.

In the resulting practical implementations of all these goals – I will actively engage the film and Television students at Arcada as assistants and production resources, thus introducing more production scenarios into their education.

2. RESEARCH QUESTION AND WORKFLOW

My primary research question is: *how can visual knowledge building be utilized as a paradigm encompassing the development and assessment of a model for augmented local television?*

This question can be articulated in six more specific questions or stages:

2.1 Stage 1: Ontology and epistemology

What is the form and nature of these new concepts in relation to local television – i.e. what aspects are we examining? (The ontological question).

How can e.g. social video, ubiquity, learning objects, visual knowledge building and other important concepts be understood in relation to what there is today: traditional TV and video on demand?

How can the analytical relevance of these concepts be determined ie. how can the process of DIY media production using new methods of production be operationalized in a way that allows the effects to be studied? (The epistemological question)

What are the functional expectations: what could the notion of "augmented local television" be – if it could be all it could be? I.e. what are the non-articulated forecasts of what kind of TV people want? This stage leads to:

2.2 Stage 2: Expert opinions and predictions

What do media professionals think are the main desired qualitative attributes of augmented local TV?

How do these interact with the traditional cable content?

What does an ecological perspective allow us to understand about the impact of technological change on the economical and cultural aspects of normative television?

These questions will be answered by conducting qualitative interviews in the form of focus group studies and in-depth interviews (grounded theory) with people from different qualified professional fields with regards to television, sociology, economy and ITC.

As a result of stage two I will have the specifications for the functionality of augmented local TV – which in turn allows to proceed to:

2.3 Stage 3: Constructing a model and a prototype

How can bar-codes and mobile terminals be interlinked to create an augmented version of local TV?

How should the broadcasting system be developed to allow for user generated content?

What are the packaging and metadata problems that need to be resolved?

Once there is a prototype version that can be tested I will proceed to:

2.4 Stage 4: Pilot testing and gathering of empirical data.

How do digital natives and late adaptors compare in their view of the system functionality and how it conveys information?

What kind of technical issues arise during the pilots? How can the feedback of the pilot results be used to improve on the model?

This stage will hopefully be completed with a real implementation of the system. If the parts necessary to build this technology mash up aren't available I will do the pilot test on a conceptual model and proceed to:

2.5 Stage 5: Analysing the results of the pilot test

Once the main empirical phase is completed. I will proceed to analysing the results using the methods described below.

2.6 Stage 6: SWAT analysis and conclusions

At this stage I should have the answer to the main question about how visual knowledge building is conveyed through the new media interface.

I can now conclude and discuss further work by answering the following: What are the strengths, weaknesses, opportunities and threats to this new media system?

How can the concept of augmented local TV be expanded upon?

Forecasting from a media ecology pov: how will the introduced change in technology affect the economics and cultural influences of local TV? Here media ecology will be used as an assesment tool for studying the affects of the platform.

Altogether, although these six stages present a multitude of questions – it should be feasible to answer most of them by breaking down the larger problem into these simpler subquerys

In order to improve on my results I will probably have to iterate the stages 3 and 4 a few times to develop the final version of both the conceptual and technological version of the system.

3. METHODS

I approach my research questions from an empirically holistic point of view. My principal methodology is that of pheonomenography as defined by Marton (1981).

My motivation for the choice of method is that Phenomenography seeks to reveal understanding of a phenomena from the participant's point of view rather than that of the researcher. The phenomenon is inseparably embedded in how the perceiver interacts with the environment – learns – and then reevaluates his position in the matter (Marton, 1981)

The empirical part of my work will be carried out through pheonomenographical studies including in-depth interviews, observations and as an important part – analysis of audiovisual material produced by the agents.

Because of the time perspective of my focus on Helsinki as "World design capital 2012" I will also use another qualitative and closely related method – i.e. the case study. This choice will allow me to purposefully sample the way people use the technologies developed – and assess how their learning experience develop during their contact with the system.

The gathered data will be transcribed, coded, and then analysed through (Trekreem, 2008):

- Meaning unit analysis

- Discussion of the interpretations compared to the transcriptions

- Categorization and definition of the outcome space.

- Filtering out of the results

- Presentation of the results.

4. FUNDING AND SCHEDULE

As a result of my research I aim at producing a Ph.D., in the form of an article dissertation. I could start the research process immediately if I get accepted as an doctorate student.

4.1 Funding

I am currently participating in the ICT-SHOK Flexible Services project "Social Video", which is funded by TEKES. The funding will extend into to the end of may 2010.

I'm applying in the immediate future for further full-time funding from my current institution Arcada University of Applied Sciences. Arcada has started a new fund which fits my profile perfectly and the prospects for receiving further funding are promising – however, this funding can only be claimed if I have an accepted research plan.

I will most likely take part in next phase of the Flexible services program - more specifcly in the research thread called *everyday services.* This program can mean an additional source of funding.

4.2 Schedule

Assuming I get the funding to conduct my studies as a full time student, I expect to finish my studies in the customary four years generally proposed for a Ph.D. degree.

5. ACKNOWLEDGMENTS

My warm thanks to the EUROITV2012 peers who reviewed my research plan! I have tried to take your comments into account as well as possible.

6. REFERENCES

[1] Avila, J. A. G. and Strate, M.C. L.: Understanding MEA, In Medias Res 1 (1), Fall 1999.

[2] Avila, J. A. G, Carvajal, M.: Integrated and Cross-Media Newsroom Convergence Two Models of Multimedia News Production P The Cases of NovotЋcnica and La Verdad Multimedia in Spain

[3] Convergence, May 2008; vol. 14: pp. 221 - 239.

[4] Bauman, Z.: Liquid Modernity. Cambridge: Polity,2000

[5] Bowman, S.and Willis,C. : We Media. API Media Center paper. http://www.hypergene.net/wemedia/download/we_media.pdf (accessed 7 March 2010).

[6] Dijk van, J. and and de Vos, L.: Searching for the Holy Grail: Images of Interactive Television, New Media Society, 2001.

[7] Gere, C., Digital Culture, 2003. Reaktion books Ltd.

[8] Hall, E. A System for the Notation of Proxemic Behaviour". *American Anthropologist* **65**: 1003–1026. doi:10.1525/aa.1963.65.5.02a0002 , 1963.

[9] Jenkins, H., Convergence Culture: Where Old and New Media Collide. New York University Press 2006

[10] Klinenberg, E.: Convergence: News Production in a Digital Age The ANNALS of the American Academy of Political and Social Science 2005; 597; 48

[11] Lynd, R. and Meddel Lynd, H.: Middletown. A Study in Contemporary American Culture. New York: Harcourt, Brace and Company, 1929.

[12] McLuhan, M.: Understanding Media: The Extensions of Man. New York: McGraw-Hill, 1964.

[13] Marton, F. (1981). Phenomenography: Describing conceptions of the world around us [online version]. Instructional Science, 10, 177-200. Retrieved March 11, 2010, from http://www.ped.gu.se/biorn/phgraph/misc/constr/phegraph.html

[14] Ong, W.: Orality and Literacy: The Technologizing of the Word. London, New York: Methuen, 1982

[15] Postman, N.: Amusing Ourselves to Death: Public Discourse in the Age of Show Business. New York: Viking, 1985.

[16] Sonvilla-Weiss (ed), S.: (e)Pedagogy P Visual Knowledge Building Rethinking Art and New Media in Education. Peter Lagn, 2005.

[17] Storsul, T. and Stuedahl, D.: Ambivalence towards Convergence. GЉteborg University, Nordicom 2007.

[18] Trekreem, A.: Dako 2: Forskningsmetodik Del II: Kvalitativ ansats och metodik. Ѓbo Akademi.

[19] Uljens, M. (1996). On the philosophical foundation of phenomenography. In G. Dall'Alba & B. Hasselgren (Ed.), Reflections on Phenomenography (pp. 105P130). Goteborg: Acta Universitatis Gothenburgensis.

[20] Zavoina, S. and Reichert, T. : Media Convergence/Management Change: The Evolving Workflow for Visual Journalists', Journal of Media Economics 13(2): 2000;143P51.

Upcoming Challenges for Streaming Media Content Distribution

Nafis Iskander and Artur Lugmayr
Department of Business Information Management and Logistics, Tampere University of Technology
P. O. Box 541, Korkeakoulunkatu 8
FIN-33101 Tampere, FINLAND,
+358504446542
{nafis.iskander }@tut.fi

ABSTRACT

The Internet has played a pivotal role in disseminating information to consumers globally. It offers a unique marketplace where a service-centric approach has been achieved by offering value added services to consumers. Theoretically, this does entail a shift in paradigm for media as ninety per cent of the latter's revenues are engulfed in traditional media such as TV, film and video. The key research question is the following "What roles do media stakeholders, producers, and the industry play in this evolving streaming media for content distribution?" The analysis will encompass these three criteria and develop key research propositions in the streaming media domain. The research methodology of observatory nature will be implemented using an inductive approach where 1. Interviews will be conducted of five to ten Finnish media companies, 2. Public documents (multiple secondary sources) will be accessed to develop industry trends such as Finland statistics, government periodicals, industry statistics and reports, EU publications etc and 3. Surveys conducted to research other media stakeholders comprising of government based surveys, organizations based surveys, and academic surveys under the rubric of ad hoc surveys. Empirical research will also be undertaken where performance tests will conducted which will formulate a data collection process against a backdrop a media economic framework to be developed in the course of this research.

Keywords

Internet, Streaming Media, Media Producers, Media Industry and Media Stakeholders

1. INTRODUCTION

The Internet has been the norm of practices consumers apply in their everyday life. The Internet is the source of a plethora of activities ranging from information acquisition, entertainment services, electronic commerce etc. Consumers choose to use the Internet to service their daily needs such as banking, shopping, education, entertainment, healthcare etc. Internet penetration has been certified by the increase of consumer access over the past decade in both US and Finland. The Federal Communication Commission statistics in the US shows consumer access over a number of technologies have witnessed a phenomenal growth for Internet. For instance, in the US, the number of ADSL subscribers grew from a 5.1 million from 2002 to 29.9 million in 2008 [1]. In Finland, the number of broadband subscriptions grew from 0 to 2.1 million from 2001 to 2008 [2]. This growth provides a glimpse that Internet usage is growing and is not likely to wither over the next decade. Scholarly interest lies into the kinds of informative sources absorbed by the Internet usage. Therefore, the research is in the upcoming trends of media content distribution for streaming media and the roles of the primary actors in this domain i.e. the stakeholders, the industry and the producers in shaping the latter.

1.1 Research Objectives and Issues

To recapitulate the research question posed at the beginning "What roles do media stakeholders, producers and industry place in this evolving streaming media for content distribution", the objectives hinges upon three core actors of the media domain. Therefore, it is imperative to initially seek out the possibilities of interactions between these players in this new challenge. The following objectives provide a more specific glimpse of the research dilemma

- To explore media production process and business value chains in the age of web 3.0
- To evaluate production processes and distribution models of streaming media
- To build a media economic framework of streaming media content distribution
- To apply the media economic model and generate performance tests

These objectives evolve from certain research questions that dominate the streaming media domain. The following section chooses to address these questions and hypothesis.

1.2 Research Questions and Hypothesis

The hypothesis stated below provides an experimental and evidentiary natural hypothesis that requires proof of empirical and observatory methodologies to ascertain its accuracy.

In the 21st century, the revival of traditional media is dependent upon the survival and growth of streaming media. Given the nature of this hypothesis, the research questions in lieu with the research dilemma and objectives poses careful scrutiny under the realm of this research:

- Is traditional media eroding? Is streaming media next generation?
- What constraints dictate production process in traditional media?

- What constraints dictate production process in streaming media?
- What is the benchmark and tools used to evaluate business entities?
- What kinds of audiences play a role in shaping media in this age?
- Which metrics for audience measurement are required for advertising?
- How does advertising change, and which potential do streaming media provide?

Given these research questions, it will be prudent now to examine the methodologies deployed to conduct the research.

1.3 Research Methodologies

The research methodology of observatory nature will be implemented using an inductive approach of data collection for the three criteria that play a key role i.e. media producers, industry and media stakeholders, where 1. Interviews will be conducted of five to ten Finnish media companies, 2. Public documents (multiple secondary sources) will be accessed to develop industry trends such as Finland statistics, government periodicals, industry statistics and reports, EU publications etc and 3. Surveys conducted to research other media stakeholders comprising of government based surveys, organizations based surveys, and academic surveys under the rubric of ad hoc surveys. Empirical research will also be undertaken where performance tests will conducted which will formulate a data collection process against a backdrop a media economic framework to be developed in the course of this research. Section 2 will provide the research perspective of the three key players in this domain. It will be wise to provide a theoretical perspective of the streaming media content as the next section discusses the latter.

1.4 Theoretical perspective of Streaming Media Content Distribution

Streaming Media is the type of media that incorporates audio, video and text in real time and distributes it over the Internet. The streaming media content where the keyword content is the information that could be audio, video or text. Technological innovations have made it possible for this kind of media to gain notoriety over the Internet by promoting value added services. These technological innovations include online platforms, open source online platforms, novel hardware and software, delivery channels, content delivery networks, integration of existing solutions for voice and video with streaming media etc. The value added services could be podcasts, video on demand, online video, online gaming, iptv, voice over ip, email, multimedia messaging services, short messaging service etc. The dilemma occurs when traditional media choose to assimilate into streaming media to provide value to its consumers. The assimilation process failed to deliver state of the art products and services. Elements of traditional media i.e. newsprint, TV channels and media syndicates undergo a constant hurdle to produce more content with lesser resources in the shortest possible time to compete with streaming media. Traditional media companies have been quite slow in mitigating to this explosion of the Internet and its adaptation in the Internet Age

For example, "In early 2000, the music industry had seen stable CD sales growth for several years. However, as online access to free content through Napster grew in popularity, the music industry was too slow to change. CD sales slipped and at its peak, Napster traded 15 billion songs. This bruised industry addressed the experiences consumers were demanding. The release of iPod proved that new technology could actually grow revenue, and in 2004 the music industry saw significant growth again." Since the traditional media confront such challenges, some notable statistics needs to be highlighted as a thinking cap for the former to consider into its future strategies. For instance: a Jupiter Research survey of more than 4,000 people revealed that consumers spend more hours a week online than watching TV [10]. Internet TV is an enlightening illustration of streaming media because Internet TV allows viewers to choose the video programming they want to watch from thousands of channels from literally all over the world [11] whereas traditional media such as TV has limitations in this aspect.

Given the theoretical shift in trends from traditional to streaming media in content distribution, the next three sections focus on the relevance and the research perspectives the three criteria/players bring to the trends in streaming media content distribution.

2. MEDIA PRODUCERS

Electronic media production hinges upon the core principle of its ability to broadcast itself to its audience. In order for that to materialize, a broadcast station is definitely a prerequisite. A broadcast station is comprised of multiple functions. It is an advertising medium, an entertainment medium, an information medium, and a service medium [3]. However, in a broadcasting environment, content producers base their content production upon the responsibilities towards its other stakeholders i.e. consumers and advertisers. However, one interesting observation is the relationship of these two stakeholders i.e. they are both directly proportional to the success of a producer's content. For instance, if content created by a producer is broadcasted and the number of audiences it amasses grows over a period of time, then proportionally the advertising revenue will also increment. This is illustrated by the Figure 1 below.

Figure 1 Relationship between Broadcast content, advertisers and consumers

With this background, it will be prudent to consider these roles under the scope of the Internet. The purpose would to provide a defining platform for streaming media. The interest lies in the shift of paradigm of consumer audience from regular channels of communication and entertainment to Internet based services. Unlike broadcast, cable TV, or satellite services, your eligible audience is almost 500 million people who have broadband internet service around the world [4]. This has translated into business benefits for companies engaged in marketing products or advertising domain. As the above section depicted the relationship between content, advertisers and consumers, it will be interesting to note that according to E-marketer, spending for Internet video

advertising in the US has nearly tripled in 2007 to 640 million dollars from 2006's 225 million dollars [5, 6]. Therefore the opportunities to monetize and capitalize on such a big market is not only appealing for the advertisers but for consumers also as they have variety of choice with incumbent and novel companies entering the marketplace with value additions products and services. The objectives of research for media producers will divulge upon interview based questionnaires encompassing the following goals:

- To highlight the role of markets and the dilemma of the media enterprises and managers in the Internet domain
- To provide the management perspective of ensuring thoughtful, decisive and agile management techniques in media production firms
- To identify technological trends and its impacts on the media production process and innovations
- To identify key advantages of newcomers media companies over existing incumbents in this domain
- To correlate a relationship between the media producers and its interaction with other actors in the scope of this paper identified above i.e. media industry and media stakeholders.

These interviews will be conducted with some of the known media production firms in Finland within the span of five to ten companies. Interviews will be transcribed and postulated further. Using an inductive approach, these interviews will serve to provide research propositions. Such propositions will be instrumental in developing the media economic framework envisioned in this research to conduct further empirical work. The next player to be discussed is Media Industry.

3. MEDIA INDUSTRY

Media can be defined as the technology, form, and content to communicate and interact with information, knowledge, or wisdom [12]. It centers upon two core concepts that dictate it i.e. communication and interaction. There have been array of media technologies which center upon the publication and distribution of media elements such as text, image, video that have evolved over time. Examples of these technologies range from cave paintings, book publication, web, newsprint, TV, radio, to podcast, live streaming, IP TV, video on demand, to ubiquitous computation in this age. Convergence in the media context merits to be elucidated. "Convergence is ushering in a new epoch of multimedia, in which voice, data and images are synergistically combined onto a single network to render better, more efficient and more innovative services to users [13]". Television or TV came into existence in the past century as the media object of dissemination of information and entertainment. It is constantly evolving with the advent of digitalization, end-user experiences, end-user preferences, program content, and societal perception. Convergence in this domain materialized with the digitization of broadcasting entities where growth of the TV became phenomenal. However, such convergence also leads to creation of new media types and technologies such as hybrid media (IPTV, it, Web TV etc).

Such challenges have had some successes in the convergence of print media with digital TV. Driven by the merger of broadcasting and IT/computing technologies, traditional TV becomes a multimedia platform accumulating more and more different functionalities and usage options for its end users. For example, content transmission over heterogeneous networks and additional end-device functionalities also provides end users with the option to use and retrieve TV programs differently—for example, to use content (inter)actively via a backchannel. The Internet is such a valuable resources that it creates the environment and platform for the consumers to adapt, consume, absorb and relate content. Users who upload and share various types of files on websites often tag them with keywords to make them easier for themselves and others to find [14]. The UK based Telegraph newspaper sensed such convergence of its print media with the TV and decided to embark upon the migratory nature of broadcasting its content online since 1994. Even though it commenced with a micro-site approach for launching Telegraph TV on-demand video content available online, the news media did evolve into the next generation platform to realize the full potential of online video. The Telegraph leveraged Bright cove's next generation online video platform to help drive their next generation video strategy to ensure flexibility, efficiency, improved user engagement and the company's extensive markup language(BEML) [9]. Such actions did bring about dynamic change for the Telegraph where after implementation, commendable outcome pointed to increased traffic (consumer), video plays, and better advertising revenue(click through rates)

The objectives of the media industry provide the following core set of goals:

- To report a chronological set of innovations and changes in the media industry over the past century
- To provide the current patterns and trends in media industry
- To observe the paradigm shift in patterns and research key clusters dictating these shifts
- To provide a vision and outlook of the present and future of the media industry in light of the challenges posed
- To provide an academic overview of the research potentials associated with the changing paradigm in the media industry

To summarize the above narration presents a thorough background of the media entity and an evolution of the industry. Publications, government periodicals, industry periodicals and organizations relative secondary sources will be acquired and studied to develop research trends. These trends will be factored into the media economic framework to be designed. The next and final criteria to be researched is the Media Stakeholders

4. MEDIA STAKEHOLDERS

Media Stakeholders that invest in providing streaming media content distribution services are spread across a broad range of business functions and capabilities. These stakeholders have mushroomed over the past decade over the growing popularity of Internet. It has been established according to an online publishers association study, more than 140 million people (69%) have watched video online with 50 million (24%) doing so weekly [7]. Given these statistics, innovation is the key towards gaining a leading edge in the Internet marketplace. Content delivery is no longer dictated by the producing attractive content for audience rather it is governed by audience tastes during media content production. These stakeholders usually invest in streaming media ventures on the basis of three criteria that play an influential role

in their investment and returns. These three are technological innovation, business performance, and customer's base. The stakeholders as in any media industry control the distribution rights of the content distribution and play a distinctive role in the entire process of the content creation and distribution n the streaming media domain. For instance, there are six phases that encompass the streaming media distribution process i.e. content creation, content management, content delivery, content publication, reporting/analytics and finally monetization. The stakeholders control their investment in these phases ensuring and taking into account the risks involved in each phase. This relationship helps the producers provide content that enthralls the stakeholders and reduces investment costs. Based on this the following specific set of goals specifies the objective:

- To provide a modeling approach of the stakeholders investment and returns in this segment
- To report evaluation metrics and inference of the metrics to provide a core understanding of the stakeholders perspective
- To research and study the different phases of the streaming media distribution process and correlate its impact on the stakeholders investment
- |To correlate a relationship between the industry, stakeholders and producers

Surveys comprising of government surveys, ad hoc surveys, organization and academic survey will dominate the core of this research to provide the relevant data for media stakeholders. This data will then evolve using inductive approach into research criteria which will factor into the media economic model in this research**.**

5. FUTURE WORK

Future work will envision the inductive approach of data collection by using secondary sources such as interviews, questionnaires, ad-hoc surveys and research publications, periodicals and government documents etc. The interviews will undergo visits to media industry in Finland with respect to five to ten companies in this domain. Interviews will comprise of questionnaires, tape recording sessions, and downloading tools to transcribe recording sessions into transcripts, interview annotation and compilation. The companies will be classified into different research and target groups stemming from academia, to professionals etc. Surveys and documents research will encompass library visits and research institutions that conduct surveys based on the subject at hand. These sources will then be formulated and research propositions developed from them to create a media economic framework. |The realm of this framework will be realized by evaluation metrics and inference of the latter. The timeline for the tasks are highlighted below

Year 1 Establishment of State Of Art and Evaluation

Year 2 Media Economic Framework Design and Creation

Year 3 Performance Tests Empirical Research

Year 4 Evaluation of Performance test, Data Collection

6. REFERENCES

[1] Ding, W. and Marchionini, G. 1997 A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.

http://fjallfoss.fcc.gov/edocs_public/attachmatch/DOC-284932A1.pdf

[2] Finland Statistics.

http://www.stat.fi/til/sutivi/index_en.html

[3] Peter K Pringle, Michael F. Starr 2006. 5th Edition, Electronic Media Management. Focal Press, ISBN Number ,978-0-240-80872-7, 12-15.

[4] Mavrogeanes, R. 2008, Video Strteaming as a Service, White paper published in Streaming Media magazine, Innovation Series 5.

http://www.streamingmedia.com

[5] Schumacher-Rasmussen, E., 2008, Best Pracices For Content Monetization Through Online Advertising, White paper in streaming media magazine, Innovation Series 5

[6] McKelvie, A, Picard, Robert G, 2008, The Growth and Development of Young Media Firms, Published Journal of Media Business Studies, Routledge, 5(1): 1-8.

[7] Vividas, 2008, Best Practices in Media, Broadcast and Entertainment, Simplify, Engage and Monetize, White paper published in streaming media magazine.

[8] BrightCove, 2008, Trends Shaping Online Video in 2009 and Beyond. White paper published in streaming media magazine.

[9] http://www.businessinsider.com/brightcove-ceo-company-profitable-cash-flow-positive-2009-6

[10] "Tapping Active TV Viewers Leads to Tremendous Opportunities", Maven Networks, Innovation Series 8, Streaming Media white paper

[11] "The Promise of Open Source Video for the Enterprise", Kaltura, Innovation Series 8, Streaming Media white paper

[12] Artur Lugmayr, "Ambient Media", Journal Novatica, vol. 33, pp 35-39, July-August 2007

[13] Shawn De Freitas1, Begoña Bagur University of Brigthon, Di8it LTD, "A case study analysis on digital convergent design: SkyNet platform" TISCP Adjunct Proceedings of EuroITV 2007." vol. 35: TICSP - Tampere International Centre for Signal Processing, 2007, p. 3

[14] Ed Hi Cho, Palo Alto Research Center, "The Social Web: Research and Opportunities," IEEE Computer Society, Computer Magazine, Volume 41, Number 9, September 2008, p 88

INDUSTRIAL CASE STUDIES

Blending CGI with real 3D

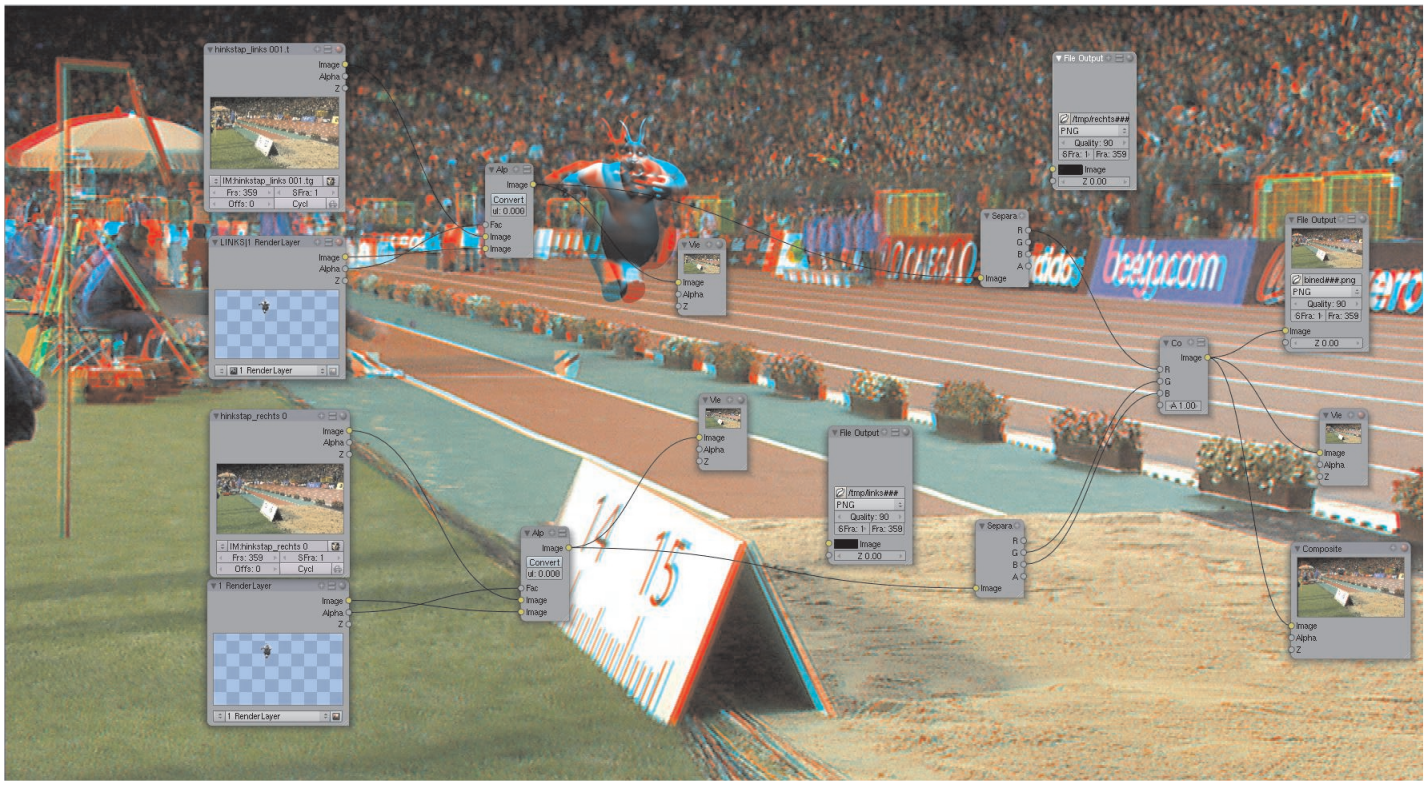
VRT-medialab

At VRT-medialab we carry out research into the creation, management and distribution of media content. Our focus is on the Flemish media market, but VRT-medialab pursues collaboration on a European and worldwide level.
VRT-medialab is the technological research department of the VRT, the public service broadcaster of Flanders, Belgium.
http://www.vrtmedialab.be

About the author

As an engineer-architect Kasper Jordaens uses his creative-technical background to follow up on fast changing technology and to implement these in new media concepts.

kasper.jordaens@vrt.be

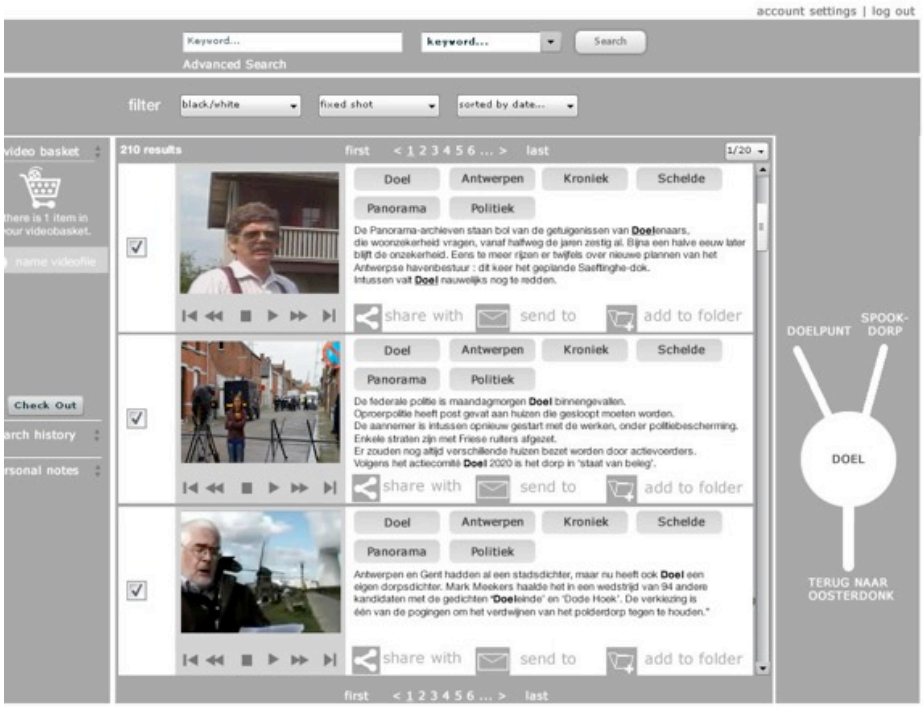


Integrating computer generated images with camera footage is not a trivial task. Camera parameters have to be correct throughout the scene. Not only lens parameters but also camera motion has to be tracked. When repeating this exercise for 3D footage this means you have to track an extra camera and the position of the cameras on the rig. If you don't have an exact match (interocular distance, focal length, back dimensions, ...) for the real and the virtual cameras you're in for a serious headache.
This Presentation "Blending CGI with real 3D footage" will give an in-depth technical overview on the complete process:
With all the Memorial Vandamme 3D footage material we had for 3DTV testing I just couldn't resist tampering with it to add something to that third dimension. Analyzing and tracking the 3D camera footage led to an animated model allowing me to start integrating and placing virtual objects in accordance with the depth script resulting in a complete and balanced compelling 3D experience.
It was an interesting excercise of testing the boundaries of what is possible, tracking problems, finding solutions and comparing those to existing solutions.

VRT Media lab Redesign of VRT video archive search

Currently, the search appliance for exploring and annotating the video archive of the VRT (Flemish National Broadcaster) is not exactly an intuitive, user-friendly experience. Based on OpenText, the web interface does not meet the rich, interactive web experience that modern users expect from a search application. Nevertheless the application is used on a daily basis by TV researchers, journalists and archivists.

In order to bring the search appliance into the 21st century VRT-medialab committed an extensive Usability research on the current tools in order to define a major design overhaul with the focus on usability, smart information architecture and rich, interactive design. We have applied the Contextual Inquiry method for observing 14 people (8 researchers and 6 archivists) in their naturally occurring context. As an outcome we defined 3 personas and crafted a detailed task analysis for each persona. Based on these results we redefined the user interface design. The



redesign takes into account rich metadata representations such as face recognition and geographic location. Within the new design, users can personalize their search interface, track their history and keep personal notes for each research subject. Query results can be shared with other researchers for better collaborative exploration. Users can drill down through search results by means of filters and a semantic wonderwheel. Individual video clips can be trimmed in a rough cut editor and added to a video cart for easy exporting to a high definition file or tape.

About VRT-medialab

At VRT-medialab we carry out research into the creation, management and distribution of media content. Our focus is on the Flemish media market, but VRT-medialab pursues collaboration on a European and worldwide level. VRT-medialab is the technological research department of the VRT, the public service broadcaster of Flanders, Belgium.

What is #dekoers?

#**dekoers** is an interactive dashboard for live pro cycling events. As rich TV interactivity is hard to realize on a TV screen we believe that this interactivity will move to a secondary screen where simple, widget-style applications will provide a more immersive TV experience. Be it on an tablet-like device, a laptop or a smartphone, #dekoers combines TV broadcasting with a synchronous web-based application overlay where users can enjoy realtime cyclist tracking, twitter integration, realtime infographics and live videostreaming.

#dekoers also offers participants the possibility to re-experience the event and watch a highlights summary, based on the twitter activity at one particular moment.

#dekoers is made with web standards technologies: Javascript, HTML5, SVG and runs on any HTML5 compliant web browser. It is currently being tested at major flemish cycling events and should be ready for primetime for The Tour de France 2010.

About VRT-medialab

At VRT-medialab we carry out research into the creation, management and distribution of media content. Our focus is on the Flemish media market, but VRT-medialab pursues collaboration on a European and worldwide level. VRT-medialab is the technological research department of the VRT, the public service broadcaster of Flanders, Belgium.

About Hendrik Dacquin

For more than 12 years, Hendrik Dacquin has been researching and prototyping new interactive experiences with Alcatel-Lucent and currently with VRT-medialab.

Study on user engagement and interaction with overlay applications for ITV

Angélica Ñáñez
Barcelona Media – Center of Innovation
Av. Diagonal 177, 9th floor, 08018 Barcelona
+34 932381400
angelica.nanez@barcelonamedia.org

Raquel Navarro-Prieto
Barcelona Media – Center of Innovation
Av. Diagonal 177, 9th floor, 08018 Barcelona
+34 932381400
raquel.navarro@barcelonamedia.org

1. HIGHLIGTHS

We hope to inspire attendees to our presentation in creating new interactive applications or services for ITV.

The highlights of our talk include:

- A summary of some cognitive and emotional aspects explored in order to explain user engagement and satisfaction with overlay interactive sport application (Figure 1) In particular, we focus on the relationship between highly motivating content for the user and their perceptions of service usefulness.

- A look of the impact of personal differences in the user interaction, attending aspects such as motivation, user age and previous experience with interactive applications or services for TV.

- A look at our early framework for interaction and user engagement with ITV. Our results allowed us to answer some early queries and to identify some key aspects to taking into account in order to improve the user engagement with interactive applications for TV. These key aspects are a part of our framework draft that will provide practical guidance to industry players in the creation of new interactive content and services for ITV.

Figure 1

2. BIO OF THE AUTHORS:

Angélica Ñáñez obtained her degree on Psychology from the Central University of Venezuela, specialised in Clinical Psychology. She is working in her doctoral research on Social Psychology at the Autonomous University of Barcelona. She is currently working as a researcher in the User Interaction Lab at Barcelona Media-Innovation Centre, where she has participated in several research projects in the area of Human Computer Interaction (European funded projects and projects with private companies). In particular, she is working in several projects concerning IPTV, Internet on mobile and social networks.

Dr. Raquel Navarro-Prieto is currently the head of the User Interaction Lab at Barcelona Media Innovation Centre where she leads multiple research projects in the area of Human Computer Interaction: 10 national and European funded projects as well as several projects with companies such as HP, Alcatel-Lucent, Havas Media Group and Mediapro among others. In addition, she currently lectures at Universitat Pompeu Fabra (Barcelona). Previously, she coordinated all research work at the Interaction Lab of Universitat Oberta de Catalunya, Barcelona, for two years; worked 3 years at Motorola Research Lab (UK) and has industrial experience from Apple Inc. (USA), and HP (Spain). Raquel obtained her PhD on Cognitive Psychology. In addition to writing numerous papers, she has been part of and chaired several program committees for international conferences and has worked as advisor for the EU Commission.

3. SHORT COMPANY BIO

Barcelona Media Innovation Centre is a technology centre dedicated to applied research on communications and the media, and to the transfer of the resulting knowledge and technologies to the industry.

New Paradigm Models in Digital Media Business (IP TV and Mobile TV)

Zvezdan Vukanovic
Assistant Professor
Faculty of International Economics, Finance and Business
and
Faculty of Information Systems and Technology
University of Donja Gorica
20 000 Podgorica, Montenegro
Telephone # +382 67 813 126
E mail: zvezdanv@t-com.me

Abstract

The purpose of this article is to analyze the role of the economies of aggregation in new and digital media (IP TV and Mobile TV) business. The author argues it is necessary for international corporations to provide a holistic response that regards economics as a set of mutually interactive aggregate segment in order to position global media industry and businesses in highly competitive and volatile markets. Therefore, the article proposes the creation and adoption of four strategic approaches in emphasizing the economies of aggregation: the Triple- and Quadruple-Play Bundling Strategies, Two-Sided markets, Complementors and Network Externalities. Adding and implementing these complex microeconomic approaches will result in accelerating and increasing the competitiveness and innovation, market share, demand, profit of global media companies. The author empirically proves the validity for IPTV and Mobile TV market growth by calculating the projected number of worldwide users from 2009 to 2015 based on the analysis of nineteen leading international research and consulting agencies.

Keywords: Digital media, interactive media, media economics, Economies of aggregation, Triple- and Quadruple-Play Bundling Strategies, Two-Sided Markets and Network Externalities, Complementors, IP TV and Cellular–Mobile TV.

Introduction: The Role of the Economies of Aggregation in Media Business and Industry

Due to the increasing global competition economists generally agree that an important feature of any modern macroeconomic theory is an explicit aggregation of the microeconomic behavior of all agents in the economy. During the last two decades, user-generated content, mass customization, and personalization have replaced mass production in the media business. Accordingly, economies of scale have been substituted largely by economies of scope. Meanwhile, global hyper-competition has fragmented niche markets and stimulated media corporations to search for more original and innovative services. To be profitable, innovative services in the media industry and business have to be diffused and distributed efficiently and effectively via various cross-media platforms (broadband, multicast, and convergent digital models). Ideally, the economies of aggregation can be based on four crucial strategic and economic concepts: Triple- and Quadruple-Play Bundling Strategies, Two-Sided markets, Complementors and Network Externalities. It is important to point out that all four concepts are fundamental in terms of increasing the demand and diffusion of innovation.

The Economies of Aggregation in Global Media Business

The importance of the economies of aggregation in global media business is particularly important as present and future markets are in the process of global expansion. The potential for global competition will increase by about 300% between 2007 and 2037. It is important to notice that in terms of the average return on invested capital, the media industry (1963–2003), together with pharmaceuticals, household and personal products, and computer software and services, represents the most profitable global industry [1].

The increased importance of the media industry has grown incessantly over the last 15 years as a result of a continual deregulation of broadcasting industries as well as new digital convergence. Its increased importance is particularly reflected in economic respects. The global media industry encompasses over $1 trillion [2] and accounted for about 22% of the total information industries' revenue [3]. In its annual media forecast, Pricewaterhouse-Coopers (2008) projects that global media revenue will grow by an average of 6.6% per year, reaching $1.29 trillion in 2012.

Empirical Evidence for Global Growth of Mobile TV and IPTV Markets

The author empirically proves the validity for IPTV and Mobile TV market growth by calculating the projected number of worldwide users from 2009 to 2015 based on the analysis of nineteen leading international research and consulting agencies dominantly based in the USA, UK, China and India. The international consulting agencies that calculate the projected number of worldwide users of Mobile TV in the period 2009 – 2015 include In Stat Research, Datamonitor Research, Inofonetics Research, RNCOS, Informa Telecoms and Media, Juniper Research and McKinsey Research. Their projections of worldwide Mobile TV users show an increase of Mobile TV subscribers from 2009 to 2015 for approximately 242% (78.5 to 190 millions).

On the other hand, the research covering the projection of worldwide IP TV subscribers was based on the reports of the following international consulting agencies MRG – Multimedia Research Group, Gartner Research, Infonetics Research, Strategy Analytics, Open IPTV Forum, IMS research, Parks and Associates, i Supply Corporation, RNCOS, Pyramid Research, Canalys, Companies and Markets. Their research shows an increase of IPTV subscribers from 2009 to 2013 for 202% (42.2 to 85.5 millions).

Benefits of Bundling Strategies

The bundling of services is defined as marketing two or more components of the same service together as a package at a special price [4]. Bundling is the strategy of adopting, competing, and differentiating corporate portfolios in volatile media markets. Bundling strategies add value to different services by inventing economical packages that are convenient to use. As such, bundling can create economies of aggregation for information goods if their marginal costs are very low, even in the absence of network externalities, economies of scale, or economies of scope. Bundling stimulates multi-product media firms to innovate. The profitability of bundling results from economies of scale in the tied market [6]. Competitive pressures and changing consumption habits are encouraging media firms to market bundles of services that include television, telephony, and Internet access [5]. If implemented efficiently bundling strategies may mitigate competition by inducing more differentiation [7, 8, 9], enhance profit; ensure customer loyalty (thereby decreasing customer churn); and increase consumer choices, market shares, and average revenues per user. Moreover, bundling helps reducing churn and protects against incursions from new competitors.

A quadruple and triple-play bundling strategy implies the utilization of multiple services, devices, and technological domains (TV, broadband, telephony, and mobile telephony), but one network, one vendor, and one bill.

A triple and quadruple play bundling strategies are very common in the software business (e.g., bundle a word processor, a spreadsheet, and a database into a single office suite) and in the USA and EU cable television industry.

It is advisable to point out that triple-play bundling strategy is most successful when there are economies of scale in production and economies of scope in distribution, marginal costs of bundling are low, production set-up costs are high, customer acquisition costs are high, and consumers appreciate the resulting simplification of the purchase decision and benefit from the joint performance of the combined product.

The Role of Network Externalities in the Media Industry and Business

Network externalities were originally introduced in the communications network literature. Before the invention of telecommunications, Internet, and digital media, the effect of network externalities was less visible and dominant.

The existence of network externalities is the key reason for the importance, growth, and profitability of global media industry in the new, digital, and network economy. Unlike in many other businesses, in the media services industry the benefit from consuming increases with the number of other people consuming [10].

Historically, indirect network externalities have influenced the outcome of technology competition in many markets, including AM stereo, color television, videocassette recorders, CD players, laser disc players, and personal computers [11, 12, 13]. More recently, as analog technologies give way to digital technologies that require new software, indirect network externalities will play an important role in the evolution of a wide range of technology markets [14].

An extra subscriber to the media network brings additional benefits to current subscribers. Similarly, the loss of a subscriber reduces benefits to current subscribers. For example, a telephone is of little value if no one else is using it, of moderate value if only a few of one's potential contacts use it, and indispensable if everyone uses it. Obviously, the value of consuming a certain TV channel by only a few consumers has increased with the number of other subscribers. Economists refer to this phenomenon as network externalities. Accordingly, a product or service possesses network externalities if the utility one derives from it is a positive function of the number of other people who consume it.

In their fundamental study on the network economy, Shapiro and Varian described the rules that guide the dynamics of networks. They argued that it is necessary to achieve a critical mass in the network to grant positive feedback. They also explored the effects that a network is subject to such as network externalities and lock in. Network externalities and critical mass are considered crucial aspects when taking into account the whole network with its multiple stakeholders such as partners, customers, consumers, shareholders, employees, investors, regulatory sectors, governments, and so on [15].

However, media management and economics researchers in the mobile-TV and IP TV industries have been slow to respond to the growing importance of network economies and externalities in new product and service adoption. For instance, most new product models in the management science literature assume that new products are autonomous and that the adoption of new products is not affected by the presence or absence of complementary products [15]. These assumptions are being called into question in almost every durable product market in the network economy, where firms rarely act alone to create new products, and products rarely function in isolation [16].

Features of Two-Sided Markets

Digital media services operate platforms that must attract two sides of a market that exhibit network externalities in order to succeed [17]. A market is said to be two-sided if at any point in time there are: (a) two distinct groups of customers; (b) the value obtained by one kind of customers increases with the number of the other kind of customers; (c) an intermediary is necessary for internalizing the externalities created by one group for the other group [18]. Accordingly, "a two-sided market is a market that requires a platform for different groups to interact and exhibits network externalities such that more participants from one group will encourage additional participants from the other group [17]. In a two-sided market, two groups interact through an intermediary, or platform, that accounts for the externalities between the groups. In the media context, the platform is the broadcast company (or companies) and the two interacting groups are advertisers and viewers.

As the media industry sells a joint product to two different categories of buyers: the medium itself to advertisers, and the medium content to media consumers (readers, TV-watchers, web-surfers, etc.) the media firms thereby operate in two different industries and get their profits from both.

Importance of Complementors

Complementors is a term used to describe businesses that directly sell a product (or products) or service (or services) that complement the product or service of another company by adding value to mutual customers; for example, Intel and Microsoft (Pentium processors and Windows), or Microsoft & McAfee (Microsoft Windows & McAfee anti-virus).

Moreover, a platform becomes more attractive to consumers as the number of its complementors increases [19]. In media industry, platform competition results in a richer structure of interactions: the number of complementors in each platform affects pricing (and therefore profits) of both platforms and all complementors [19]. Naturally, it also affects utility of all consumers.

In the two-sided markets the value of a platform for one side of the market increases with the number of players in the other side of the market that adhere to it. Therefore, the author argues that platforms and complementors always benefit from an increase in the number of complementors in their same platform.

Brandenburger and Nalebuff in their influential book 'Co-opetition' point out that the more complements there are and the closer their relationship to the products supplied by the industry the greater the potential profit within the industry [20]. Complementors therefore have the ability to raise barriers to entry if incumbent firms have already developed products that are compatible with the complementors' [21].

Advantages of Network Externalities in Media Markets (IP TV and Cellular–Mobile TV)

Unlike with traditional cable and satellite TV, the global media market is not saturated with products and services relating to IP TV and cellular–mobile TV. As such, these two types of media are more effectively positioned to utilize the benefits of network externalities.

In addition, the main advantage of IP TV, Internet, and cellular–mobile TV over other media such as newspapers, radio, and TV is that they are not one-way, but rather two-way, mediums of communication. Their content can be efficiently accessed, distributed, as well as customized and repurposed to suit individual consumers' needs and preferences. Because of the fact that it can be efficiently accessed, Internet takes advantage of "locational monopolies," which represents the monopoly from being physically close to the customers [22].

Strategic and competitive advantage of IPTV

There are at least twelve key components in which IPTV outperforms other rival TV platforms (digital terrestrial, cable or satellite). These segments include: 1. Better internet infrastructure; 2. Increase in broadband speeds; 3. Better compression technology; 4. High broadband penetration world-wide; 5. Increased high quality content; 6. Positive support from major telcos; 7. Better interactive applications; 8. Advances in video-on-demand storage capabilities; 9. The development of Secure Conditional Access/DRM solutions for IPTV; 10. Next generation, intelligent IPTV set top boxes; 11. Better bundling and media diversification strategies; 12. Effective application of the ''triple play'' and a two-sided market structure; 13. recommendation, better program guides and greater personalization/tailored (targeted) advertising [23].

The Main Benefits of Mobile TV

The main benefits that will improve the potential growth of mobile TV industry include: strong brand infrastructure, targeted services, changing market landscape, upgrading of 3.9G/4G networks [24], new and niche content formats, targeted advertising, better suitability for interactive applications, personalized content programming and presentation, as well as improved customer relations. Also, it is advisable to point out that effective application of mobile TV benefits from "long tail" economics through access to niche content as well as leverages their existing brands through new delivery (web based technologies) channels and content formats relations [25].

Barriers to the use of Mobile TV

The main barriers facing a more rapid development of the mobile TV are:

- Conflicting broadcast mobile TV standards: ISDB-T (Japan); DVB-H (Europe and US), DMBX (Korea) and MediaFLO (US), 1seg, MBMS, TDtv, CMMB, TMMB, DMB-T/H, CDMB, CMB and ATSC-M/H.
- Short battery life
- Low image quality
- Lack of content
- Consumer disinterest
- Government regulation

Global trends in Mobile TV – Further Research

In order to fully leverage the concepts of dual markets and network externalities, mobile TV and IPTV industry leaders might consider option of converging digitally their content production and services. The model of digital convergence between Mobile TV and IPTV business should be dominantly based on the following factors:

- convergence of networking interface and interoperability path
- integration of services and re-purposing of content distribution
- timely facilitation of content transfer from mobile TV to IPTV
- common program production and formatting
- cross-platform profile access and billing
- cross-platform content access
- PC access to IPTV and MTV (Mobile TV)
- Transferring program from IPTV to mobile TV.

Concluding Remarks

In this paper the author proposes a new paradigm shift in digital media economy by adopting the role of the economies of aggregation. As such, the author argues that in order to position effectively global media industry and business in highly competitive and volatile market, it is necessary for international corporations to provide a holistic response that regards economics

as a set of mutually interactive aggregate segments. As such, it is advisable to adopt innovative approach in emphasizing the economies of aggregation that feature the Triple- and Quadruple-Play Bundling Strategies, Two-Sided markets, Complementors and Network Externalities.

The author empirically proves the validity for IPTV and Mobile TV market growth by calculating the projected number of worldwide users from 2009 to 2015 based on the analysis of nineteen leading international research and consulting agencies (In Stat Research, Datamonitor Research, Inofonetics Research, RNCOS, Informa Telecoms and Media, Juniper Research, McKinsey Research, etc). The figures show an increase of Mobile TV subscribers from 2009 to 2015 for approximately 242% (78.5 to 190 millions). On the other hand, the research shows an increase of IPTV subscribers from 2009 to 2013 for 202% (42.2 to 85.5 millions).

In addition, the author argues that the digital switchover, the untapped market of new broadband and mobile/cellular technologies, combination of near ubiquitous broadband access, increasing digital convergence, lower equipment costs and home networking, competitive pressure among cable, satellite and telecom operators, consumer familiarity with a 'pull medium' – cellular phone and the Internet, all point to a rapidly developing market for both IPTV and Mobile TV.

Accordingly, the network externalities and two-sided markets provide international media companies with the opportunity to deliver bundled triple play services with a common infrastructure and a common user experience in order to maximize revenue and maintain customer loyalty. The effect of Two-Sided markets and network externalities will help media companies to attain the economies of scale needed to leverage the creative content and technology.

In addition, the digital convergence of mobile TV and IPTV will extend positive effects of network externalities and provide international media companies with valuable models of content re-purposing and related media diversification. This is particularly true if we take into consideration that only 1.8% of all mobile phone users worldwide that are presently estimated at 3.3 billion are Mobile TV subscribers. On the other hand, only 2.2% of all TV users worldwide (currently estimated at 1 billion) are IPTV subscribers.

REFERENCES

[1] Grant, R. M., (2008). *Contemporary Strategy Analysis*. Oxford, England: Blackwell.

[2] Vizjak, A., & Ringlstetter, M. (2003). *Media management: Leveraging content for profitable growth*. New York: Springer.

[3] Compaine, B., & Gomery, D. (2000). *Who owns the media?* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

[4] Shaw, M. (2006). *Commerce and the digital economy*. Armonk, NY: Sharpe.

[5] Baranes, E. (2006). *Bundling and collusion in communications markets* (Working Paper No. 06–17). NET Institute.

[6] Whinston, M. D. (1990). Tying, foreclosure and exclusion. *American Economic Review*, *80*, 837–859.

[7] Carbajo, J., de Meza, D., & Seidmann, D. J. (1990). A strategic motivation for commodity bundling. *Journal of Industrial Economics*, *38*, 283–298.

[8] Chen, Y. (1997). Equilibrium product bundling. *Journal of Business*, 85–103.

[9] Seidmann D. J. (1991). Bundling as a facilitating device: A reinterpretation of leverage theory. *Economica*, *58*, 491–499.

[10] Hoskins, C., McFadyen, S., & Finn, A. (2004). *Media economics: Applying economics to new and traditional media*. Beverly Hills: Sage.

[21] Ducey, R. V., & Fratrik, M. R. (1989). Broadcasting industry response to new technologies. *Journal of Media Economics*, *2*, 67–86.

[12] Farrell, J., & Shapiro, C. (1992). *Standard setting in high-definition television*. In M. N. Baily & C. Winston (Eds.), *Brookings papers on economic activity*. Washington, DC: Brookings Institute.

[13] McGahan, A., Vadasz, L. L., & Yoffie, D. B. (1997). Creating value and setting standards: The lessons of consumer electronics for personal digital assistants. In D. B. Yoffie (Ed.), *Competing in the age of digital convergence*. Boston: HBS Press.

[14] Yoffie, D. B. (1996). Competing in the age of digital convergence. *California Management Review*, *38*(4), 31–53.

[15] Foss, N. J., Kristensen, T., & Wilke, R. (2004). Corporate communication in the emerging network economy: A provider of common knowledge. *Corporate Communications: An International Journal*, *9*, 43–49.

[16] Shapiro, C., & Varian, H. R. (1998). *Information rules: A strategic guide to the network economy*. Cambridge, MA: Harvard Business School Press Books.

[17] Mahajan, V., Muller, E., & Bass, F. M. (1993). New-product diffusion models. In J. Eliashberg & G. L. Lilien (Eds.), *Handbooks in OR and MS* (Vol. 5, pp. 349-408). Amsterdam: North-Holland.

[18] Cortade, T. A Strategic Guide on Two-Sided Markets Applied to the ISP Market, *Communications & Strategies*, no. 61, 1st quarter 2006, p. 17.

[19] Carrillo, J. D. and Tan, G. Platform Competition: The Role of Multi-homing and Complementors, Working Paper #06-30 October 2006, The Networks, Electronic Commerce, and Telecommunications ("NET") Institute.

[20] Grant, R. M., (2002). *Contemporary Strategy Analysis*. Fifth Edition, Wiley, John & Sons, Incorporated.

[21] Haberberg, A., and Rieple, A., (2008). *Strategic Management*, Oxford University Press.

[22] Jansen, D. W. (2006). *The new economy and beyond*. Cheltenham, England: Edward Elgar Publishing, Incorporated.

[23] Ashun, P. (2008). *IPTV Overview*. Future Media & Technology: Keeping the BBC relevant in the digital world. BBC. Presentation.

[24] Korhonen, J. *Mobile TV – A New battle of Standards*, Cantlab Wireless Limited, 2009, Cambridge, U.K.

[25] Mobile TV UMTSF/GSMA Joint Work Group, Final White Paper "*Mobile TV: The Groundbreaking Dimension*", Retrieved January 15, 2007, from http://www.umts-forum.org/servlet/dycon/ztumts/umts/Live/en/umts/MultiMedia_PDFs_Papers_joint-mobile-tv-group-white-paper

From Reach to Engage: Harnessing the Marketing Potential of Interactive TV

Ana C. Andrés del Valle
Accenture Technology Labs
449 rte des Crêtes
Sophia Antipolis – 06902 France
ana.c.andresdelvalle@accenture.com

OVERVIEW

Until recently, TV was considered the most powerful marketing tool available. However, technology evolution and changes in consumer habits have since made the Internet the number-one advertising sector, ahead of TV, in some markets (such as the UK & Denmark). The loss of TV's marketing supremacy is also a consequence of the digital transformation that most businesses are undergoing. By digitalizing the access to their customers, the best brands do not only reach their consumers, but also spend their energy engaging them in experiences that delight them. Traditional TV has lost its place in a world where experiences, not messages and ads, build businesses and brands. So, what can iTV actually offer marketers?

iTV is developing fast, and is gradually being adopted as an emerging technology with the disruptive potential to blur the lines between physical environments and digital technology. It provides business with an opportunity to engage with consumers anywhere and anytime in totally new and differentiated ways. iTV will be increasingly relevant to brands as marketing professionals start considering it:

- ***a key piece of the Digital Transformation ecosystem*.** iTV is a renewed customer access platform that, unlike the Web, has a privileged link to the home environment. Televisions will be coupled with other home digital systems, from where their Internet connectivity will make them the natural gateway to outside the home. However, iTV is currently suffering from a larger technology fragmentation than the Web or the Mobile worlds, and proliferation of TV widgets, different flavors of STB, home systems, etc. will make marketers hesitant to bet on a single approach. To succeed, iTV providers should offer marketers the means and the tools to link home consumer experiences with what is happening elsewhere: in stores, on the street, on the Web…
- ***a high-quality audiovisual and interaction platform on potentially any device***. iTV providers need to evolve from focusing on a PCTV-like approach to delivering audiovisual media on mobile devices and developing new channels into consumer spaces where the projection of highly-targeted messages in specific places, allied with relevant interaction, will give brands more "persuasion power". New sources of marketing revenue will come from unique consumer experiences.

Accenture Technology Labs is analysing and experimenting to create examples of how interactive TV could become part of the new "digital marketing journeys" that businesses want to offer their customers. In parallel, the Labs also assesses business models that could help iTV providers to successfully offer attractive marketing solutions. As Accenture participates in the digital transformation of many companies, the Technology Labs works towards creating a full Consumer Integrated Digital Experience for our clients' customers.

AUTHOR'S BIO

Dr. Ana C. Andrés del Valle is a researcher at the Accenture Technology Labs, in Sophia Antipolis, France. She has over 10 years experience in the area of Human Computer Interaction for Telecommunications. She is currently driving the TV SENTINEL (TV Services ENhanced Through INternet and game Experience Levers) at the Accenture Integrated Digital Experience group. In the past, she worked for AT&T Research and collaborated with France Telecom R&D.

ABOUT ACCENTURE

Accenture is a Global management consulting, technology services and outsourcing company, helping its clients become high performance businesses and governments. With approximately 177,000 employees and clients in more than 120 countries, it served 96 of the Fortune Global 100 and more than three-quarters of the Fortune Global 500 (2008). Named to Fortune's Top 50 list of the world's most admired companies, it has also ranked no. 1 in CRM consulting, according to CRM magazine.

User experience of next generation content delivery: users are developing a new sense for quality

Marjan de Jonge
TNO Information and Communication Technology
Eemsgolaan 3, 9727 DW Groningen
The Netherlands
+31 505857743
marjan.dejonge@tno.nl

Oskar van Deventer
TNO Information and Communication Technology
Brassersplein 2, 2612 CT Delft
The Netherlands
+31 152857078
oskar.vandeventer@tno.nl

Pieter Nooren
TNO Information and Communication Technology
Brassersplein 2, 2612 CT Delft
The Netherlands
+31 152857207
pieter.nooren@tno.nl

ABSTRACT
This paper presents preliminary results of user experience research for video-on-demand (VoD) services. The purpose of this study is to identify the aspects that determine the quality of experience (QoE) for the user and to research alternative options for service providers to optimize the use of these resources, given that resources are limited.

The novel Eureka Celtic Rubens network-aware service architecture enables service providers to tune various aspects of the user's VoD service experience. A group of 15 test users took part in the service experience test of the Rubens demonstrator. Based on in-depth interviews with these users, we draw the following preliminary conclusions:

- Users are developing a new sense for quality when more choice in quality in media content is offered.
- The QoE for video-on-demand services is multi-dimensional. Lead time and choice in content are found to be similarly relevant to the user as the technical quality (e.g. standard or high definition).
- Paying for premium quality touches all these three QoE dimensions.

These insights help service providers to optimize the QoE for their users over networks that do not have unlimited resources.

General Terms
Design, Human Factors.

Keywords
Quality of Experience, user experience, usability, user interaction, content delivery, video on demand.

1. INTRODUCTION
The increasing consumption of on-demand, personalized media puts the traditional media content delivery chain under pressure as resources such as bandwidth and storage space are scarce, and need to be optimized. For distributors, typically cable and telecom operators, one of the major challenges is: how to deal with the increased consumption of personalized and long tail content (e.g. Internet TV) and the ever increasing triple-play bandwidth requirements (e.g. transition to HDTV) in an economically sound way?

The hypothesis of the Eureka Celtic Rubens consortium[1][1] is that a high perceived quality of experience (QoE) can be achieved for video-on-demand (VoD) services without requiring costly absolute bandwidth guarantees nor unlimited content storage space. A set of demonstrators has been built to show the feasibility of a novel network-aware service architecture and to validate the hypothesis with test users.

This paper presents preliminary results with regard to a user evaluation on the QoE, usability and perception aspects. The paper is organized as follows. Section 2 introduces the Rubens project and the three dimensions of QoE. Section 3 focuses on the research method. Section 4 presents the results of the user experience research. Section 5 summarizes the conclusions.

2. Rubens: optimizing the QoE
QoE relates to how user perceive the quality of an application or service. To capture such a subjective measure, either by subjective tests or via objective tools is quite a challenge. Techniques for video and (mobile) TV have been investigated [2, 7, 9, 10]. Our qualitative research method will be further described in Section 3.

The Rubens network-aware service architecture [12] brings two solution directions to address the challenge of optimizing and measuring QoE (in the absence of absolute bandwidth guarantees). First, the architecture provides a more efficient way to deliver media content through an intelligent combination of transport and storage to lower the costs of media delivery.

[1] The Rubens project is a European project within the Eureka Celtic framework. It has the following partners: Alcatel-Lucent, British Telecom, Deutsche Telekom AG, France Telecom R&D, IBBT, KPN, Telefónica I+D, TNO ICT, University of Zaragoza. This work has been partly supported by Agentschap NL, an agency of the Dutch Ministry of Economic Affairs.

Second, the architecture allows for optimization of the QoE. The QoE approach in Rubens focuses on three dimensions[11].

1) **Video quality**, determined by, among others, the resolution (high versus standard definition) and (the absence of) hick-ups and artefacts in the image. Its impact on the underlying delivery network is relatively straightforward: a higher resolution requires more bandwidth in the network.

2) **Choice in content**, ranging from popular content to all content on the Internet. Simply put, the more content a consumer can access, the better his experience will be. This dimension has a clear impact on delivery networks that employ caching or storage: content that is watched by many users can be efficiently delivered from caches in the network, thus saving bandwidth in the network.

3) **Lead time**, which is the time that elapses between the moment a consumer selects a piece of content and the moment to watch it. In general, QoE will degrade as the waiting time increases [8]. However, the idea behind the lead dimension is that for certain types of content, e.g. movies, people plan in advance when they want to be able to watch and do not mind that the content is not available right away. During lead time, the media delivery network has more flexibility: it can finish other media delivery requests before starting the delivery of the new one, or it can already start delivering the new one to a buffer, thus freeing up network resources later in time.

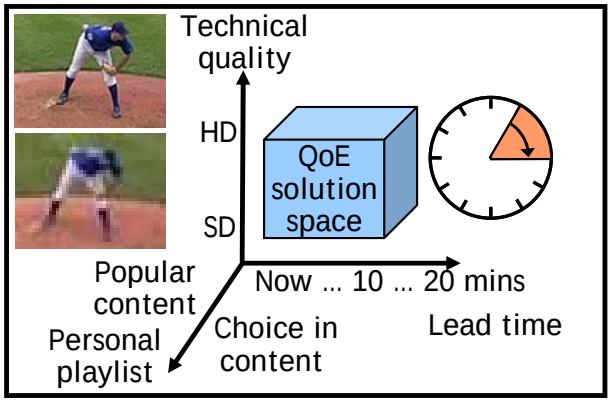


Figure 1: Three dimensions of QoE for VoD services

In the Rubens approach, the three dimensions are considered in combination. For example, a consumer wants to view an High Definition (HD) movie. Because of network congestions, the movie cannot be delivered right away in HD. Instead of simply stating "request denied" as is done today, the consumer is offered a number of options, such as: view the HD movie in 10 minutes, watch the movie now in Standard Definition (SD), or watch another (more popular) movie in HD now. Thus, Rubens aims to find the optimal user experience.

3. Method: qualitative investigation

3.1 Research questions

The user experience research focused on the evaluation of the following research questions:

1. How important is quality for the user, e.g. difference between SD-HD?
2. How does the user choose for quality? High quality content versus lower quality content? What drives the user to make this choice?
3. How does the user react to decreasing quality or interruption while using a service?
4. Preferences with regard to quality and conditions, e.g. waiting for better quality, pay more, choose other high quality content?
5. What are user needs with regard to information design (in communicating with application or even other users)?
6. Presentation of content: What format/concept offers optimal perception of quality and usability?

3.2 Approach: user-centered design

To perform the research, the user-centered design (UCD) approach was used [5]. In UCD, designers engage actively with end-users to gather insights that drive design from the earliest stages of product and service development, right through the design process. UCD consists of a wide range of methods and tools in order to identify user needs, perceptions and the usability of a product. An overview of methods used in the UCD process for the development of an interactive TV concept can be found in [4,6]. Different methods are often combined, ranging from qualitative methods such as focus groups, interviews, observations, probes, workshops and personas to quantitative methods such as log files and surveys [3].

In order to obtain in-depth insights in QoE and needs with regard to quality and content, service perception and information design of the Rubens VoD services qualitative research was performed. Users were interviewed about their experience and perceptions for about 90 minutes. To be able to perform the interviews based on the set of use cases enabled by the Rubens architecture, a visual presentation of a user interface (mock-up) was built (see 3.4).

3.3 Use cases: congestion, now what?

The user experience research was based on a set of use cases representing relevant situations where the Rubens architecture enables service providers to optimize QoE for end users.

3.3.1 Single user using one single application

A user wants to watch a video from a VoD service. When the required resources are not available now, the system notifies the user and offers an alternative, e.g. lower technical quality or longer lead time.

3.3.2 Two users with the same application

Two users each want to watch a video from a VoD service. When insufficient resources are available, the system offers interaction with the two users to optimize their use of resources.

3.3.3 Two users with different applications

This use case is similar to the previous one, but the applications are different. One user wants to watch a real-time video stream while the other user starts up a game.

3.4 Mock-up: experiencing Rubens QoE

The mock-up simulates a HD video on demand user environment where users can choose from different categories of content, view movie information (contents, IMDB rating, quality information). Furthermore, the mock-up focuses on different information and feedback flows presented to the user in case of high network load and quality degradations allowing for different scenarios to be tested. More details on the mock-up are available from [12].

Figure 2: Mock-up of Rubens user interface

3.5 User characteristics

Respondents were selected to fit in three target groups: students, young professionals and professionals each with their own age requirements (see table 1). Further target group requirements were: high internet experience, more than average interest in movies, and gaming affinity.

Table 1: Overview of demographics

Number of respondents	15
Number of males	12
Number of females	3
Category “Students” (20-25)	6
Category “Young Professionals” (26-35)	5
Category “Professionals” (36-55)	4

4. RESULTS

In this section we will present the analysed results of the user experience test.

4.1 Quality matters but in different ways

The technical quality of the content (e.g. difference SD-HD) is an important driver for most respondents. Most users (80%) look for and are interested in the best quality content only (e.g. HD). SD as an alternative is only considered when at a reasonable level.

In case of a commercial service (e.g. paying for content), all participants expect high technical quality and a guaranteed quality. To be able to get the best possible quality most customers do not mind buffering (as they are used to time-consuming downloading) if this offers a guarantee for quality (see 4.3). Users owning HD equipment especially appreciate HD, they specifically search for HD content, they are prepared to pay for it, and they expect technical quality guarantees in return. The occasion is an important driver in choosing HD: watching HD with friends “cinema style” is a preferred to going a real cinema.

4.2 Choice for quality driven by mix of contextual and situational factors

During tasks, the choice of respondents for HD versus SD depends on a mix of contextual and situational factors, such as the user’s mindset (expect to get an on demand HD service and want to watch immediately), the type of movie and its importance to the user (latest movies should be in HD, a TV series can be is SD), the importance of the occasion (e.g. a movie night with friends), the mental preparation (‘I want to see this movie now in HD because I have been looking forward to it.’), the type of equipment owned (LCD TV, beamer or other advanced movie viewing equipment) and the information provided (see 4.4). Furthermore, the actual perceived quality of SD is mentioned as a crucial factor (see 4.3).

4.3 Predictability and control of quality is crucial

Most respondents indicate a need for insight in the exact quality when choosing between HD and SD (clear information on the available quality options, resource status and pre experience of the exact quality via preview or trailer). For the delivery time, users find “when I want” more important than “now”. In addition, respondents stress they want to choose both quality and time of delivery in advance (e.g. before the start of the actual viewing of the content).

All participants oppose to an interruption of the flow of the movie; quality degradations during a movie are perceived as very disturbing. Most participants are only willing to accept one interruption per movie (maximum of one small break of 5-10 minutes). If a change in quality (HD to SD, buffering) is necessary, then preferably at a time indicated by the user.

When offered the choice to view a particular movie in SD immediately or in HD later (e.g. in 15 minutes to allow for buffering), most users indicate they would prefer to wait for the movie to buffer to be able to see HD. However, as a consequence in advance buffering or pausing a movie damages the “on demand” perception resulting in users no longer willing to pay as much.

4.4 Information should be non-intrusive

In case of decreasing quality due to resource shortage, most respondents indicate they value information provided by the application or network. They prefer a clear problem description: why a quality problem occurs, how long it will last and what they can do about it. Ideally a warning should be given in advance. Information should be non-intrusive (e.g. small expandable icon/popup) to be expanded on the initiative of the user.

Alternatively, users strongly indicate that they do not want to be informed about or interact digitally in any form with other users about resources (especially those outside their own homes) in case of resource shortage. Many participants prefer the network to maintain an acceptable level of quality for all users involved.

4.5 GUI/mock-up evaluated positively

Most respondents were able to navigate the mock-up easily. Most liked features were usability factors like: simplicity of the design, ease of use, learnability, easy navigation and category selection. User feedback and information in the mock-up was designed in layers, presenting general information first with the option to activate additional more detailed information in layered windows (so details do not appear immediately but only when the users activate it). In case of degrading quality a small popup icon is presented in the lower left corner of the user interface (for more details see [12]).

5. CONCLUSIONS

In investigating the user experience of the novel Rubens network-aware service architecture, we focused on QoE, usability and perception dimensions. In this section we present the conclusions drawn from the results presented in section 4.

Our main conclusions are the following.

- Users are developing a new sense for quality (determined by QoE and usability dimensions).
- QoE, usability and perception are determined by content quality (SD/HD), lead time and choice in content.
  - The availability of different levels of quality results in more awareness of quality. Quality is thereby becoming a selection criterion for consumers. In addition, paying is a demand for quality.
  - User control is key and alternative service lead time options (e.g.buffering) to receive higher quality are linked to strict conditions (see fig.4).
  - Choice in content is a major driver in determining the value of the video service for users.
  - Information, feedback and predictability are crucial to user experience (e.g. in case of alternative service quality being informed in advance and in detail).
- Paying for premium quality touches all three QoE dimensions.

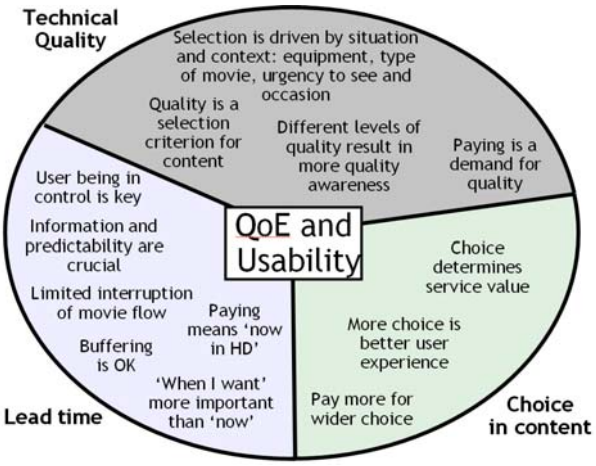


Figure 4: Sub conclusions based on three dimensions of QoE

6. Further research

Based on the results of the preliminary results of the Rubens user experience research further research could be performed. For example a more extensive pilot with a larger base of users (e.g. households) in a real operator context incorporating the mock-up and the Rubens demonstrator to validate the previous results. Furthermore additional insights and issues gathered during the course of the project could be tested (e.g. personalization, delay of content delivery).

7. References

[1] Berghe, Van den, S., Re-thinking the Use of Broadband Access for Experience Optimized Networks and Services, NEM Summit, Saint-Malo, France, 28-30 September 2009.

[2] Chipchase, Y.J., and Yanqing, C., Personal Television: a qualitative study of Mobile TV users in South Korea. MobileHCI 2006.

[3] Eronen, L., Combining Quantitative and Qualitative data in user research on digital television. In Proc. of the 1st Panhellenic Conf. PC HCI, p. 51-56, 2001.

[4] Eronen, L., User Centred Research Methods for Interactive Television, In Proc. of 3rd EuroITV, pages 15-26, 2006.

[5] ISO 13407: Human-centred design process: international standard as a basis for UCD methodologies, www.iso.org/

[6] Jones, M. Video on Demand Usability Demands, www.journalism.cf.ac.uk (2007).

[7] Knoche, H., and Sasse, M.A., Getting the big picture on small screens: Quality of experience in Mobile TV. Idea group. 2007.

[8] Kooij, R.E., Klos, V.B., Ahmed, O.K. et al., Optimizing the quality of experience during channel zapping, International journal on advances in Systems and Measurements, 1942-261x vol. 2, no. 2&3, year 2009, pp. 204-213.

[9] Kuipers, F., Kooij, R.E., Vleeschauwer, de, D., Brunnström, K., Techniques for Measuring Quality of Experience, 8th Int. Conf. on Wired/Wireless Internet Communications (WWIC), June 1-3, 2010, Lulea, Sweden.

[10] Mäki, J., Finnish Mobile TV Results, Research International Finnland, 2005.

[11] Pons, C., Optimising the Management of Rich Media - A Rubens Consortium Perspective, SIP and IMS for Next Generation Telecoms, Berlin, 23-25 September 2009.

[12] Rubens Demonstrator, Deliver on-demand content with a better experience to customers, Venue: The 5th Annual Celtic Event, part of IV Conference on FP7, Valencia, Spain, 12-13 April 2010, http://www.r2sconference.eu/sideEvents_celticEvent.php and http://www.celtic-initiative.org/Events/Celtic-Event10-Valencia/welcome.asp

TUTORIALS

ESCape: Directions for the next wave of interactive television research

Konstantinos Chorianopoulos
Ionian University, Greece
choko@ionio.gr
http://www.ionio.gr/~choko

Pablo Cesar
CWI, the Netherlands
p.s.cesar@cwi.nl
http://homepages.cwi.nl/~garcia/

ABSTRACT
This course provides the foundations for the design of interactive television experiences on the Web and beyond. It regards the human and social aspects of content editing, sharing, and control with a special focus on television for the Web. In particular, the course provides the attendants with a background on the technologies, and user behavior. We present the ESCape framework, which describes interactive TV research up to this day, as well as provides a solid ground for further research. ESCape stands for content editing (E), content sharing (S) and content control (C). We propose this taxonomy so that designers of future interactive teleivion can choose their path of research, with awareness of other directions that have been explored in the past.

Categories and Subject Descriptors
H.5 INFORMATION INTERFACES AND PRESENTATION: User Interfaces

General Terms
Design, Human Factors

Keywords
Interactive Television, video, methods, user experience, interaction design.

1. OVERVIEW
The course focuses on the design of novel user experience for interactive TV applications, and research/product opportunities in the field.

The presenters will provide a framework for understanding television research. The framework is based on the basic television related activities: content editing, content sharing, and content control. Editing corresponds to the activity of producing or re-editing available content. Sharing corresponds to a meta-content activity, such as recording and talking with friends and family. Finally, Control corresponds to the selection process, 'what to watch'.

Interactive television builds on established theories of media science and interaction design with a particular focus on how people create, consume, and share rich media content, such as videos and television. As a matter of fact, the course covers a broad range of topics such as content distribution, user interface design and development, user interaction principles, user-generated content, and sociability.

Moreover, it provides an overview of the ongoing academic and industrial research on interactive mass media. For this purpose, we review products and research projects. The goal is to try to understand why research issues such as personalization, or commercial products such as the “red-button” did not fulfill the expectations, and why other solutions such as TiVo, or YouTube have been very popular.

Finally, we present topics for further research. We will explain how current research focus on the media narratives and how the end-user can affect the content they are watching. Then, we will discuss about social television and its implications in user interfaces and computer mediated communication. Finally, we will focus on ambient intelligence and how everyday objects can be used to control and manipulate content and we will have a discussion on the impact of user-generated content in the new media landscape.

In summary, this course is very suitable for the EuroITV audience because it provides an overview of the state-of-the-art, as well as methods for evolving toward the next wave of interactive television research.

2. TARGET AUDIENCE AND EXPECTED BENEFIT
This is an introductory course that is aimed at graduate students, practitioners and researchers in multimedia, web, human-computer interaction, and media management/design who wish to gain an understanding of what is the state-of-the-art and how to design new services for interactive television. Moreover, the course is suitable for educators who want to introduce a new module on ITV and want to know the basic learning materials. The course includes a discussion on research topics for doctoral students that would like to contribute to this emerging field. Finally, the course is suitable for people in the industry who want to learn more about interactive television on the web, TV, or mobile phones.

The materials for the course include lectures and videos. The lecture material (slides) will be provided to the attendees. The course builds upon an extended survey article, which will also be provided in digital format.

Cesar, P., and Chorianopoulos, K. The Evolution of TV Systems, Content, and Users Towards Interactivity, Foundations and Trends in Human-Computer Interaction (FnT-HCI), 2(4): 279-373, 2009.

3. INSTRUCTORS' BIOS

The instructors have presented similar courses at:

- ACM WWW 2009: Interactive Television and the Web, http://www2009.org/tutorials/T8-A.html
- UXTV 2008: Introduction to User Experience Design for Interactive TV, http://uxtv2008.org/program/intro-to-itv.html
- EuroITV 2008: Human-computer interaction methods
- ACM WWW 2008: Interactive Television and the Web http://www2008.org/program/program-tutorials-TA5.html
- EuroITV 2007: Introduction to iTV: State of the Art and Future Directions
- ACM Mulimedia 2006: Interactive Television and Multimedia Systems

In addition, the first instructor has taught a similar course to a group of thirty senior undergraduate students at UdK Berlin over the spring and winter semesters 2007-2008 (http://udk-itv.wikispaces.com). Two research papers on iTV have been published as a result of that course.

Konstantinos Chorianopoulos is a lecturer in the Department of Informatics, at the Ionian University (Greece). In 2002, he founded UITV.INFO (http://uitv.info), which is a newsletter and web portal for interactive television academic research resources, industry news, and events. He is the main author of more than ten journal papers and he has lectured internationally on several aspects (design, engineering, user behavior, management) of interactive TV. http://www.ionio.gr/~choko

Pablo Cesar is a researcher at the 'Centrum Wiskunde & Informatica' (the national research institute for mathematics and computer science in the Netherlands) in Amsterdam. He has (co)authored over 40 articles (conference papers and journal articles) about multimedia systems and infrastructures, media sharing, interactive media, multimedia content modelling, and user interaction. He has given tutorials about Interactive Digital Television in prestigious conferences such as ACM Multimedia and WWW Conferences. http://homepages.cwi.nl/~garcia

Columns on Last Page Should Be Made As Close As Possible to Equal Length

Tutorial: Designing and Evaluating the Sociability of Interactive Television and Online Video

David Geerts
CUO, IBBT / K.U.Leuven
Parkstraat 45 Bus 3605
3000 Leuven, Belgium
+32 16 32 31 95
david.geerts@soc.kuleuven.be

ABSTRACT
In this tutorial, we will discuss how the social uses of television have an impact on how we should design and evaluate interactive television and online video applications. We will focus on the concept of sociability, and explain how this can be evaluated using guidelines and heuristics. We will also discuss how sociability can be studied by performing user tests, and which aspects of testing are different from studying usability. Although the guidelines and user tests are especially appropriate for designing and evaluating social television systems and online social video, parts of it are also suitable for other iTV or online video applications.

Categories and Subject Descriptors
H.4.3 Information Systems Applications: Communications Applications, H.5.2 User Interfaces: Evaluation/methodology, H.5.1 Multimedia Information Systems: interactive television

General Terms
Measurement, Design, Experimentation

Keywords
Sociability, heuristics, evaluation, social television

1. INTRODUCTION
In past EuroITV conferences, as well as at uxTV2008, social TV has proven to be an important and exciting new topic of research in interactive television. In the past year, social TV is starting to make the move from academic and industrial research labs [e.g. 1, 3, 6, 7] to the consumer market. The introduction of widgets on connected television sets (e.g. by Yahoo! and Opera) opens up a whole range of possible social TV applications, and the first social video applications on the web are taking form (e.g. Watchitoo or Clipsync). As is good practice in user-centered design, evaluating these systems early and often is important to create an optimal user experience. Although several guidelines for evaluating the usability of interactive TV exist [2, 9], and heuristic evaluation as well as usability testing is a well-known and often practiced technique, for applications being used in a social context such as the social television systems and online social video applications mentioned above, evaluating only usability is not enough. Even if these applications are evaluated to improve their usability, it doesn't mean that the social interactions they are supposed to enable are well supported.

This tutorial wants to fill this gap by teaching researchers and practitioners how to design and evaluate social features of interactive television and online video. Based on his extensive experience in performing user tests of social television systems for evaluating their sociability (e.g. [4]), the presenter will explain the practical issues related to performing user tests with iTV focused on social interactions. Furthermore, he will discuss the sociability heuristics he has created based on these tests [5], as well as several other social interface guidelines, and explain how they can be used to evaluate social television systems and online social video applications, or social aspects of interactive television and online video in general.

The proposed structure of the tutorial is as follows: first, the social uses of television as documented by several media researchers [8, 10] will be shortly introduced. They will be linked with the current state of interactive television services and online video applications, including a wide range of social TV systems and social video applications on the web, and the need for designing and evaluating sociability will be explained. After this, an overview of sociability evaluation methods focused on social interaction will be discussed, including small exercises. Then, an overview of twelve sociability heuristics the presenter has developed will be given, along with an explanation of how to use them to evaluate iTV and online video. Finally, a practical hands-on session will be held in which the participants can apply the sociability heuristics to an online social video application such as Watchitoo (or – due to the fast changing nature of the area – another suitable application that will be available at the time).

2. SCHEDULE

- 1/2 hour introduction to the social uses of (interactive) television and online video
- 1 hour discussing sociability evaluation methods
- 1/2 hour overview of sociability heuristics
- 1 hour practical exercise

3. TARGET AUDIENCE

The target audience for this course are researchers and practitioners that design or evaluate interactive television or online video applications, and that want to focus on social aspects of iTV and online video. The tutorial requires no specific background knowledge, as most concepts related to the content of the tutorial will be explained in detail.

4. LEVEL OF THE TUTORIAL

The level of the tutorial is beginner.

5. TUTORIAL HISTORY

This tutorial was first organized as a workshop for interaction designers and usability professionals as part of the EU CITIZEN MEDIA project at the K.U.Leuven (22/1/2009). Based on the positive responses from participants, the instructor was invited to teach the course for the Eindhoven Birds of a Feather (BOF) group of CHI Nederland at the Technical University of Eindhoven (26/3/2009). The course was also taught at EuroITV2009 (3/6/2009). For EuroITV2010, the course will be modified from the previous edition at EuroITV2009 by broadening the focus to all video based applications, and including a broader range of evaluation methods. Given the fast changes in this area, care will be taken to include the most recent examples of social iTV and online video applications, especially the application used for the interactive exercise. Finally, minor revisions will be made based on previous participants' comments.

6. PRESENTER'S BIO

David Geerts has a master in Communication Science at the K.U.Leuven and a master in Culture and Communication at the K.U.Brussel. He was project leader of the Mediacentrum of the Katholieke Universiteit Leuven (Belgium) for several years, and now leads the Centre for User Experience Research (CUO). He is involved in several research projects on user-centered design and evaluation. Furthermore, he acts as content manager for the post academic course Human-Centered Design. David has over ten years of experience in teaching for a diverse audience: e.g. introductions to new technologies for secondary school children, practical seminars in web design for university students and in-depth courses on usability topics for practitioners. Currently he teaches a master course in Human-Computer Interaction for students in communication science, economy and industrial engineering, as well as usability design courses and workshops for practitioners. David Geerts has organized workshops and SIGs at CHI2006, CHI2007, CHI2008, EuroITV2007 and EuroITV2008 He finished his doctor's degree on Sociability of Interactive Television, for which he has developed twelve heuristics for designing and evaluating social television interfaces. David Geerts is co-founder and chair of the Belgian SIGCHI.be chapter, and was program chair of EuroITV2009, the 7th European Interactive TV Conference.

7. REFERENCES

[1] Boertjes, E., Schultz, S. and Klok, J. Pilot ConnecTV. Gebruikersonderzoek. Public report of the Freeband Project, TNO (2008)

[2] Chorianopoulos, K. User Interface Design Principles for Interactive Television Applications. In International Journal of Human-Computer Interaction, 24, 6, Taylor Francis (2008)

[3] Coppens, T., Trappeniers, L. & Godon, M. AmigoTV: towards a social TV experience. In Masthoff, J. Griffiths, R. & Pemberton, L. (Eds.) Proc. EuroITV2003, University of Brighton, (2004)

[4] Geerts, D. 2006. Comparing voice chat and text chat in a communication tool for interactive television. In Proc. NordiCHI 2006, 461–464.

[5] Geerts, D., & Grooff, D. D. (2009). Supporting the social uses of television: sociability heuristics for social tv. In Proceedings of the 27th international conference on Human factors in computing systems (pp. 595-604). Boston, MA

[6] Harboe, G., Massey, N., Metcalf, C., Wheatley, D., & Romano, G. The uses of social television. In ACM Computers in Entertainment (CIE), 6, 1 (2008)

[7] Harrison, C., Amento, B., 2007. CollaboraTV: Using asynchronous communication to make TV social again, In Adjunct Proceedings of EuroITV2007, 218-222.

[8] Lee, B. & Lee, R.S. How and why people watch tv: implications for the future of interactive television. In: Journal of advertising research, 35, 6 (1995), 9-18

[9] Lu, Karyn Y. Interaction Design Principles for Interactive Television. Master's Thesis, Georgia Institute of Technology (2005)

[10] Lull, J. The social uses of television. In: Human Communication Research, 6,:3, Spring, (1980) 197-209

Tutorial: Experience prototyping layered ITV application

Rich Radka
Node – Innovation for Growth
Almogavers 125
08018 Barcelona, Spain
+34 629 731 868
richradka@gmail.com

ABSTRACT
The purpose of the tutorial is to practice methods that can be used to develop scenario-driven prototypes to reveal how users will interact with applications that are layered over the primary video content on ITV platforms.

Categories and Subject Descriptors
H.5.2 [User Interfaces]

General Terms
Design, Experimentation, Human Factors

Keywords
Rapid prototyping, open innovation, personas, modes of use, layered activities, shifting focus, multiple users' control

1. INTRODUCTION
By 2012, TV manufacturers plan to release at least 100 million televisions into the global market that will be internet-enabled and upon which applications will run concurrent with video content. Many of the attendees at EuroITV 2010 – whether they represent manufacturers, carriers, services, or enabling technologies – have a vested interest in trying to understand how consumers will interact with this growing world of applications on the TV. All too often, product and service development, especially in the telecoms space, tends to begin with technologies and/or business plans and not with the actual needs and behaviors of consumers in the real world. This tutorial aims to leverage the cross-functional expertise that will be in attendance at EuroITV 2010 in order to develop user-centered rapid prototypes of interactive experiences to share ideas about innovative yet realistic paths to pursue in ITV.

2. TUTORIAL OBJECTIVES
The layered nature of interactions will be a central characteristic of the next generation of ITV applications, but ITV will still be the TV and next generation devices and their locations in homes will continue to be optimized to view traditional or new forms of video content.

Newly created applications will exist within this context, so the most successful will stretch existing metaphors but will not discard the conventions of TV use today. (To further define the scope of this tutorial, we will primarily address fixed "big screen" experiences rather than laptop viewing, iPads, or other contexts.)

This tutorial will present an opportunity for participants to explore an holistic approach to developing interactive experiences for next generation ITV. The primary objective will be to introduce a general methodology and a set of specific tools in order to create prototyped experiences that consider existing and emerging user needs and interaction expectations while also fitting within the feasibility of technological and business realities. Participants will get the opportunity to try out the methodology and specific tools to create some application concepts. Secondarily, we will place this prototyping methodology within a larger experience-design context.

We will use the following important characteristics of layered applications to build and validate our prototypes:

1. Access to applications (always on, alerts, etc.)
2. Managing availability to social networked apps
3. Context switching
4. Account and identity management
5. Managing multi-tasking
6. Ability to share content and objects
7. Individual vs. group activities
8. Control and capture (RC, cameras, Wi-Fi/Bluetooth)
9. Integration across devices (mobiles, PCs, storage)
10. Integration with third party apps with TV (EPC, etc.)

3. TUTORIAL APPROACH
The tutorial will be highly participatory and hands-on with a bias towards doing, not just talking. Our approach will be pragmatic not academic in order to leverage the skills sets, knowledge, experience, and current real-world problems of all attendees, whether they are students, engineers, designers, or business developers. The emphasis will be on open sharing of ideas about applications, interactions, and prototyping methods in order to develop prototypes for experiences that will provide all of in the tutorial with fresh insights.

The organizers of the tutorial will provide the materials and themes for the entire tutorial. However, we also encourage participants to bring any of their own real-world application ideas, unmet user needs, and other existing business or technical challenges & opportunities to the tutorial to augment our prepared materials.

4. TUTORIAL STRUCTURE
The tutorial will follow the outline below. Specific approach for each activity will be tailored to the number of participants, level of familiarity with the methods and concepts of the tutorial, as well as the specific interests in the audience.

- **Introductions and sharing of expectations/needs from the tutorial** – Participants will explain their relationship to ITV application development and reasons for joining the tutorial. We will start to build a sandbox of opportunities, challenges, and concerns to draw on throughout the tutorial. We will clarify the agenda and confirm objectives for the day. (15 minutes)

- **Overview of factors and trends and emerging consumer expectations –** The facilitators will present relevant social trends and consumer behaviors to act us "brain food" for creating content during the tutorial. Participants may add their own perspectives, but we will reframe all ideas within a consumer-centered context to set our focus for the day. (10 minutes)
- **Explanation of scenario-driven methodology –** We will present the overarching approach, the steps within the approach and some potential tools that can be used. We will provide an argument to support our proposed approach to application development, but we will not delve deep into design thinking or social science theory. Our goal will be to only provide the minimum necessary conceptual overview so that everyone has enough context to actively practice and assess the methodology. (20 minutes)
- **Introduction of "problem area" briefs for each group –** The audience will break into small groups and the facilitators will provide different situation briefs for each group. Each brief will be broadly thematic and will provide enough breadth to be used for the rest of the tutorial. (10 minutes)
- **Groups use the methodology to develop responses to the briefs –** Groups will engage using the methodology and specific tools that have been introduced. The facilitators will circulate among the groups to provide context, answer question and provide guidance as needed. (40 minutes)
- **Review of initial opportunities against user specification criteria –** Each group will present back initial thoughts with an emphasis on the key elements from a consumer perspective. Other participants will provide constructive critiques based on the criteria we have established and insights developed within their own group's interactions. (15 minutes)
- **Break** (15 minutes)
- **Revision of concepts and choosing a single concept for prototype development –** Back in small groups, we will transform the opportunities for the problem areas into specific user scenarios that will highlight one or more applications and/or sets of features that clearly meet particular user needs and desires (50 minutes)
- **Review of scenarios and application prototypes –** Each group will again report their progress and insights both for the sharing of interesting ideas and to receive critical feedback. (30 minutes)
- **Review process, context and next steps for further information –** We will discuss what potential next steps should be – these may vary by group – and put the day's work back into the larger context of user-centered ITV experience development. (15 minutes)
- **Final conclusions and feedback on session -** As a group, we will share key themes that have emerged from the session about process, opportunity areas, and development of user-focused ITV applications. Facilitators will elicit feedback about the form and content of the tutorial. (15 minutes)

5. EXPECTED OUTCOMES

Upon completion of the tutorial, participants will understand how to apply user-centered design to the development of ITV applications. They will also come out armed with a clear understanding about important considerations in the creation of user experiences dependent on layering applications over video (or applications that interact with other applications). The content and activities of this tutorial will provide a stimulus for new thinking, and tools for creating and evaluating applications that put the user in the center of the development process.

6. TUTORIAL LEVEL

Introductory – no prior knowledge of experience design is required to participate in and benefit from the tutorial.

7. TUTORIAL LENGTH

Half-day

8. TARGET AUDIENCE

This tutorial is relevant to all functional areas; anyone with an interest in systematic but exploratory forms of innovation, understanding the use of human-centered scenarios to develop solutions, or in framing successful experience creation.

9. RESUMES OF PRESENTERS

Rich Radka is a partner at Node, Innovation for Growth, based in Barcelona, where he delivers innovation strategies, concepts new products and services, and designs user experiences for his clients across industries and platforms. He has worked with carriers, service providers, manufacturers and start-ups in the ITV space. He has over 15 years of consulting and media production experience and received his BFA in Integrated Media from Arizona State University, in Tempe, Arizona, USA.

Recommender Systems for Interactive TV

Half-day tutorial

Paolo Cremonesi
Politecnico di Milano, Milano, Italy
p.zza Leonardo da Vinci, 32 - 20133 Milano, Italy
Phone: +39 022399-3517
paolo.cremonesi@polimi.it

Roberto Turrin
Neptuny, Milano, Italy
via Durando, 10 - 20158 Milano, Italy
Phone: +39 022399-7115
roberto.turrin@neptuny.com

ABSTRACT

Recommender systems alleviate the problem of searching for interesting content in the iTV domain where users are overwhelmed by the abundance of TV channels, programs and videos. The technologies enabling a personalized viewing experience through TV recommendations are evolving rapidly. In this tutorial participants will learn about the newest developments in the domain of recommender systems and personalized TV. The focus will be on state-of-the-art technologies that can be used to identify users' taste by using either social-based, collaborative approaches or content-based analysis of TV programs. The tutorial will also provide a deep understanding of the role played by user-based assessment methods and of the effectiveness of automatic quality metrics (accuracy, errors, coverage, novelty, serendipity). During the tutorial, participants will use an on-line TV recommender systems in order to experiment with the features of different recommendation techniques.

Keywords

Recommender systems, collaborative filtering, targeted advertising, personalization

1. INTRODUCTION

Interactive Television (iTV), differently from conventional television, allows providers (i) to track user activity and (ii) to personalize the content transmitted to the users. These tasks are accomplished by *recommender systems*, whose goal is to filter information from a large dataset - e.g., thousands of channels and movies offered by an Internet TV operator - and to recommend to the users only the content that is likely of interest and attraction to them. Recommender systems play an important role in iTV applications that are characterized by a very large amount of data. The abundance of choice users need to face with may become a burden, and the presence of a huge number of items dramatically reduces the visibility of each one, potentially inhibiting users from finding interesting TV contents. From the provider's point of view, a large catalog is expensive to be maintained, due to the cost of the multimedia material itself, of the required storage space, and of the hardware infrastructure used to stream videos from the content provider to the users. Recently, many iTV recommender systems are currently being developed in academic as well as in corporate research labs, especially after the enormous resonance of the competition organized by the American movie rental provider Netflix. Differently from traditional e-commerce domains (e.g., , iTunes, Last.fm, Amazon), recommender systems for iTV provide new challenges (e.g., real-time and scalability requirements, difficulties in collecting user ratings, difficulties in collecting TV content meta–data).

The goal of this tutorial is to give the participants an overview of how to design and evaluate recommender systems for interactive television.

The tutorial will be organized as a half-day event structured into four parts. In the first part of the tutorial a comprehensive general background on recommender systems will be provided. The integration between iTV services and recommender systems will be also explained. In the second part of the tutorial a detailed overview of both statistical and user–based evaluation methodologies will be given, along with an explanation of how to use them to evaluate iTV recommender systems. In the third part of the tutorial a case study on the development and deployment of a recommender system within Fastweb, one of the largest IPTV providers in Europe, will be presented. A deep analysis highlighting the lift in VoD sales and the changes in the users' watching behavior will be also presented. Finally, a practical hands-on session will be held in which the participants will experiment with an on–line iTV recommender system.

2. SCHEDULE

Classification of recommender systems (1 hour)

The first part of the tutorial gives insights into state-of-the-art recommender techniques and their classification. In further detail, participant will learn about *content* and *collaborative* recommender techniques.

- *Content* recommender techniques find similarity between TV contents on the basis of their metadata (e.g., genre, director, actors, plot) and recommend TV contents similar to the ones the user liked in the past

- *Collaborative* techniques predict the utility of TV contents for a particular user based on the TV contents preferred by similar users.

Evaluation of recommender systems (1 hour)

The effectiveness of a recommender system is related to the quality of its recommendations which, in turn, is defined in terms of different attributes. The second part of the tutorial will address the most important techniques and metrics that can be used to assess the quality of a recommender system.

- Relevance: ability of a recommender system to suggest TV contents that fit the user's preferences.
- Novelty: capability of a recommender system to help the user in finding surprisingly interesting TV contents that the user might have autonomously discovered but he was not aware of.
- Serendipity: capability of a recommender system to surprise the user by suggesting fortuitous and unexpected TV content that the user might not have otherwise discovered.

Case study (1/2 hour)

The third part of the tutorial will describe, as a case study, the integration of a recommender system into the production environment of Fastweb[1], one of the largest European IPTV providers. The Fastweb recommender system has been released to production environment in October 2008 and implements both collaborative and content-based techniques. The system is actually providing, on average, 30'000 recommendations per day, with peaks of almost 120 recommendations per minute during peak hours. The architecture of the Fastweb recommender system will be described and the quality of its recommendation will be evaluated by means of off-line and on-line tests.

Interactive session (1 hour)

In this last part of the tutorial a final hands-on experience with a recommender system will be set up.

- Participants will be divided into groups and will be asked to explore a TV movie catalog and to assign a set of preferences.
- A number of recommendations will be provided, each group with a different recommender technique.
- Participants' opinions about relevance and novelty of the recommendations will be collected and compared against the recommendation technique.
- Overall conclusions will summarized by the organizers

3. AUDIENCE

This tutorial is designed for a general audience interested in personalized iTV services. Its content is relevant for participants with industrial, media, technological, or service background. For the HCI community, the tutorial will present novel methodologies for navigation through media contents. For the media communities, it will provide an overview on what changes recommender systems bring to existing iTV services. For the more technical audience, this tutorial will introduce recommender systems and algorithms.

[1]Fastweb is part of the SwissCom group.

4. INSTRUCTORS

The instructors of this tutorial have been actively involved both in the performance analysis of IPTV networks as well as in the recommender system community.

Prof. Paolo Cremonesi is associated professor at the Politecnico di Milano where he teaches courses on Digital and Internet Television, Computer System Architectures, Network Performance and Reliability. He is author of several publications and book chapters concerning recommender systems, distributed and parallel computation, performance evaluation, capacity planning.

Dr. Roberto Turrin received a MSc and a PhD in Computer Engineering and he is now with the R&D department of Neptuny, as well as a lecturer at Politecnico di Milano where he teaches Digital and Internet Television and Enterprise Systems. Roberto has a number of international, scientific publications concerning performance evaluation of large-scale distributed systems and recommender systems.

5. REFERENCES

[1] J. Anselmi and P. Cremonesi. Bounding the performance of BCMP networks with load-dependent stations. In *16th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS 2008)*, pages 171–178, 2008.

[2] R. Bambini, P. Cremonesi, and R. Turrin. *Recommender Systems Handbook*, chapter A Recommender System for an IPTV Service Provider: a Real Large-Scale Production Environment. to appear, Springer, 2010.

[3] E. Campochiaro, R. Casatta, P. Cremonesi, and R. Turrin. Do metrics make recommender algorithms? In *IEEE Advanced Information Networking and Applications (AINA)*, 2009.

[4] P. Cremonesi, G. Casale, and S. Visconti. Identifying network failures and evaluating link mtbf from utilization logs. In *32nd International Computer Measurement Group Conference (CMG 2006)*, pages 703–710, 2006.

[5] P. Cremonesi, E. Lentini, M. Matteucci, and R. Turrin. An evaluation methodology for recommender systems. *4th Int. Conf. on Automated Solutions for Cross Media Content and Multi-channel Distribution*, pages 224–231, Nov 2008.

[6] P. Cremonesi and R. Turrin. Analysis of cold-start recommendations in iptv systems. In *RecSys '09: Proceedings of the 2009 ACM conference on RecommenderSystems*, pages 1–4. ACM, 2009.

[7] P. Manzoni, P. Cremonesi, and G. Serazzi. Workload models of vbr video traffic and their use in resource allocation policies. *IEEE/ACM Trans. Netw.*, 7(3):387–397, 2009.

Tutorial: Browser, Widgets and Applications: An Insight on State-Of-The-Art Technologies for the Creation of Service Frontends for Social and Interactive TV Experiences

Oliver Friedrich
Fraunhofer FOKUS
Kaiserin-Augusta-Allee 31
10589 Berlin
+49 30 3463 7145
oliver.friedrich@fokus.fraunhofer.de

ABSTRACT
The technological landscape for the generation of interactive applications enabling Social Television aspects is changing rapidly. In this tutorial participant will learn about the newest developments in this domain mostly by integrating Web content and technologies into end devices and the application development lifecycle. The focus will lie on already available technological frameworks available from the different players in the market. This includes Browser and Widget driven approaches as well as end-to-end IPTV ecosystems. This will be paired with disruptive ideas combining on the on the first look silo oriented solution with aspects from IPTV standardization as discussed inside various Standard Development Organization as the Open IPTV Forum. In this context also deployment scenarios in managed Telco networks (NGN/IMS-based) and Over-The-Top approaches will be discussed.

Keywords
Hybrid TV, IPTV, Widgets, Standardization, Reference Platforms

1. Tutorial Outline
The proposed tutorial for EuroiTV 2010 goes in line with the general scope of the conference: "web.sharing.tv.content". Having these key words in mind the following questions will be answered by the different tracks during the tutorial:

- Which technology is behind and how to develop applications for the different platforms
- How to get on these platforms being a Content Provider or Social TV Application Developer´.
- How can the presented technologies interact with each other and how do they integrate in standardized IPTV environments?
- Which potential do the presented technologies have and how are they accepted in the market.

1.1 General Overview

Beginning with a general introduction providing an overview and comparison on available technologies for interactive Social TV application development the tutorial will be structured by the different technologies itself. Each of them will be discussed showing potential limitations and deployment options in managed or unmanaged IPTV infrastructures. Each technology will be concluded by giving a short hands-on which might be a demo application and a look at the development environment and/or the source code giving a “Hello World” experience. One of the discussed technologies (potentially Microsoft Mediaroom) will be presented in higher detail including a deeper presentation of a larger application.

1.2 CE-HTML, Open IPTV Forum DAE and HbbTV

Consumer Electronics HTML (CE-HTML) has a good chance becoming a de-facto standard or smallest denominator for presenting Web content on the TV. This track will introduce the ideas behind and where CE-HTML is used:

- Philips “NetTV”: What is this about and what is possible from a technical level
- Open IPTV Forum’s Declarative Application Environment (DAE) which is CE-HTML enriched with features to allow the communication with NGN and IMS based infrastructures
- HbbTV: The CEs and Broadcasters answer to discussions on Hybrid TVs

1.3 The Yahoo! Widget Channel

The Yahoo! Widget Channel has been first on the market and integrated into various TVs already in early 2009. Incorporating Web technology (e.g. the Webkit engine) a proprietary framework

namely the Yahoo! Konfabulator engine has been adapted for TVs. Again the following questions shall be answered:

- What is this about and what is behind?
- How to built a simple Yahoo Widget!
- Deployment and Test process

1.4 Microsoft Mediaroom

This track will try to answer if Microsoft's' Mediaroom is outclassing all other technologies with regards to the development experience and of course the Quality of Experience towards the user.

- Microsoft Mediaroom and the Application Developer Program (ADK) OR How a .NET/ASP.NET developer can build an interactive IPTV application within 2 minutes and has deployed it in nearly 3.
    - The Visual Studio IDE: Building a simple applications for Mediaroom.
    - Deploying an application to the infrastructure
    - Presentation of a more complex application developed for one of our partners.

1.5 What's next? W3C Widgets and HTML5

Based on actual developments till the presentation of this tutorial in June 2010, this track can be seen as a placeholder. Current key candidates for presentation might be the following:

- W3C Widgets (W3C Widget Specification 1.0)
    - A standardized way to develop (packaged) applications for Three Screens (TV, Mobile, PC)
    - How to create a simple W3C Widgets
- HTML5
    - Rich Internet Applications on TV without proprietary plug-ins.
    - Technological benefits from HTML5

1.6 Tutorial Characterization & Target Group

The presented **half day** tutorial topics will be presented in a way which should attract anyone interested in this topic and especially the technologies behind and how these technologies could be used on upcoming Interactive and Social TV deployments. As the author is explicitly not an application developer but responsible mostly for strategic research, also decision makers are more than welcome.

Nevertheless some basic understanding on Integrated Development Engines (IDE), programming & scripting languages should exist,

Classification: Intermediate

1.7 Tutorial Presenter

Diplom-Ingenieur Oliver Friedrich is a senior research engineer at the Competence Center Future Applications and Media at the Fraunhofer Institute for Open Communication Systems (FOKUS). He holds a M. Sc. in Computer Engineering from the Technical University of Berlin, Germany with specific emphasis on telecommunication networks.

Currently he is leading the team for IPTV & Rich Media Convergence at Fraunhofer FOKUS) managing the integration of new services and working on architectural issues. In addition he is preparing his PhD in the field of personalized and interactive IPTV services delivered over Next Generation Networks.

He is actively involved different international projects and in the IPTV standardization process at the Open IPTV Forum and ETSI TISPAN. In this context he published several papers, journal articles and book chapters.

1.7.1 Tutorial History 2009

- EuroiTV 2009 Leuven, Belgium
- IEEE ICC 2009 Dresden, Germany
- Lecture on Advanced Communication System, Technical University Berlin
- IMS Workshop 2009, Berlin, Germany
- Coaching & consultancy for various Telcos, CEs, Broadcasters and technology providers

DEMONSTRATIONS

Collaborative Annotation of Videos: watching and commenting (YouTube) videos

Maria da Graça Pimentel
Roberto Fagá Jr, Vivian Motti
Universidade São Paulo
São Carlos-SP-Brazil
mgp,fagá,vmotti@icmc.usp.br

Renan Cattelan
Univ.Federal de Uberlândia
Uberlândia-MG-Brazil
renan@facom.ufu.br

Cesar Teixeira
Univ.Federal de SCarlos
São Carlos-SP-Brazil
cesar@dc.ufscar.br

ABSTRACT
In this paper, we demonstrate the CWaCTool, a social, iDTV-targeted application which allows users to synchronous and collaboratively edit multimedia annotations associated with selected video frames.

Categories and Subject Descriptors
H.5.1 [**Multimedia Information System**]

General Terms
Design, Experimentation, Human Factors.

Keywords
Collaboration, Video annotation, Social TV

1. INTRODUCTION
Capture and access applications provide a rich framework for seamlessly authoring multimedia content. In earlier work we presented the watch-and-comment (WaC) authoring paradigm to explore automated capture concepts in the context of iDTV. The WaC paradigm proposes the capture of an individual's comments in a transparent and unobtrusive way, so that corresponding annotated interactive videos are automatically generated [1].

As proof-of-concept, we implemented a prototype application, the CWaCTool [2], which supports the capture of digital ink and voice comments on individual frames and segments of the video, producing a multimedia document that specifies both the media stream structure and its synchronization. The CWaCTool also provides social and collaborative capabilities, allowing users to synchronously view each others' ink or text annotations, and discuss the annotations using a text-based chat tool, while making their own ink, text or voice annotations. In the remainder of this paper, we detail CWaCTool's implementation and use.

2. CWaCTool Demo
In this section, we detail aspects relative to annotation and collaboration in the latest version of CWaCTool.

2.1 Starting to use the tool
The latest version of CWaCTool is available at http://coca.intermidia.icmc.usp.br:6943/new.

When the application is started, users log in by typing a nickname and pressing the "Test CWaCTool" button. The main application is then started with all panels empty, if is the first run, or with cached videos and annotations, in later runs (see Figure 1).

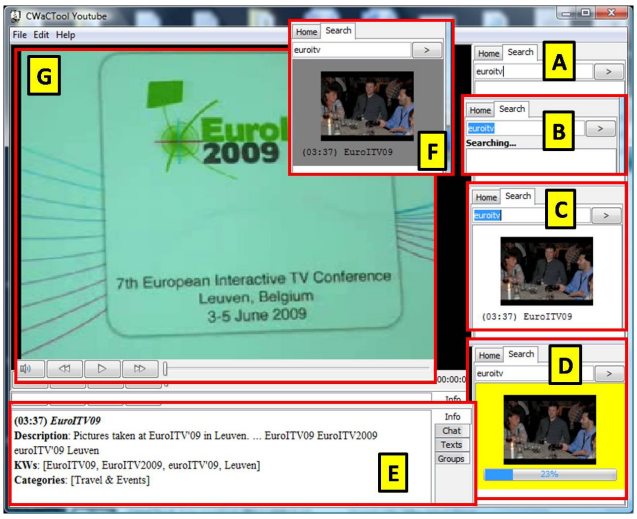


Figure 1: Searching and playing videos from *Youtube.com*

The CWaCTool has a main video canvas (Figure 1.G), a bottom panel (Figure 1.E) and a side panel to allow the user to browse current videos and search for new ones. The "Search" tab causes the given query to be processed at Youtube.com (Figure 1.A and 1.B), and the results are shown via their thumbnails, titles and descriptions (Figure 1.C and 1.E). Using a double click, users loads the video locally (Figure 1.D); the video can be played by another double click once it has been downloaded (Figure 1.F). If the user knows the URL from Youtube.com, he may use the "*File*" menu, which shows the option "*Open Youtube URL…*".

2.2 Annotation and Collaboration
Annotations can be made by any user when she is watching a video, so a video must be playing or paused. Figure 2 shows a user watching the video "Douglas Engelbart: The Mother of All Demos (1/9)", where Figure 2.A shows the video selected at the "Home" tab. The thumbnail has been created using a single click: a double click on the thumbnail causes the video to be played back in video canvas (Figure 2.B).

Regarding collaboration, and with respect to the distributed infrastructure, it is possible for several users to share a session using the CWaCTool in two ways: one is using the P2P infrastructure that allows groups to be created, as detailed elsewhere [1] and the other (not scalable) alternative is a server to which all clients connect automatically on startup (the latter is version available for download). In both cases, several collaboration alternatives are available, as detailed next.

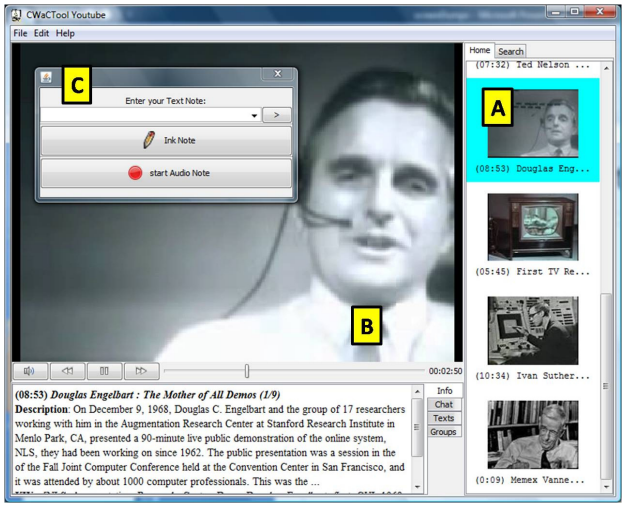


Figure 2: Video playback and annotation startup

Regarding annotation, Figure 2.C shows that the user has activated the annotation function by clicking (or tapping/touching) on the video canvas, resulting in a pop-up window to be shown with 3 options: one for typing a "Text Note", one for starting an "Ink Note", and an option for starting an "Audio Note". These annotations have distinct features relative to collaboration, as detailed next.

A "Text Note" is typed directly on the pop-up window: once the user types the note (and its "enter"), the pop-up window disappears. The notes can be seen in the "Text" tab, which shows all the notes made collaboratively by all users in a timeline relative to the video. Selecting one of these notes causes the video to play back from the time offset corresponding to the moment the user started typing the annotation. By using this feature, users can collaborate with each other asynchronously, since the sharing of the notes is automatic and transparent.

An "Audio Note" is entered by selecting the "start Audio Note" button. In this case, the button changes to "stop Audio Note" while the audio is recorded; a second click on the button stops the recording and the pop-up window disappears. In the current version, audio notes cannot be revised directly in the CWaCTool: components of the interactive multimedia document are automatically generated for each annotation session.

If the "Ink Note" option is selected, the video playback window is minimized, as shown in Figure 3. The playback continues on the smaller window, but the user may explicitly press pause, causing the playback to be suspended. The video frame at the moment of the selection is extracted and presented on the large panel (Figure 3.B), which also contains several ink-based annotation functions (shown between Figure 3.C and 3.D).

The video frames extracted by each user may be annotated collaboratively by users participating in a shared session. To share the frames relative to a given video, the user selects the icon "two users" shown on the left of Figure 3.D. When the icon is activated by one user, a pop-up window with the "Accept a session sharing" invitation is sent to all connected users. Users who accept the invitation share the frame corresponding to the particular video being watched at the time of invitation.

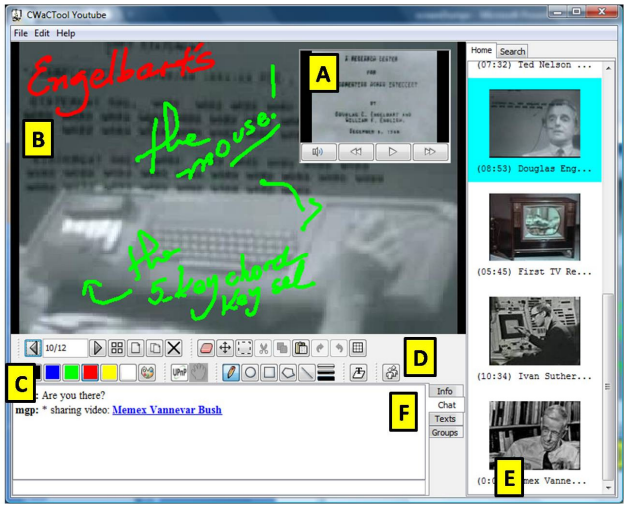


Figure 3: Ink annotations, text-based Chat and video sharing via Chat feature

In the CWaCTool, users has a text-based Chat, where they can share a video selected using the alternate mouse button (Figure 3.E). A link is shown to all users in the Chat (Figure 3.F), so only who clicks on this link will play the video.

2.3 Collaborative-based *r*ecommendation

An important collaborative feature of the CWaCTool is the video and annotation recommendation features. We implemented a simple service based on the annotations typed by the users. **W**hile the user types a new annotation, recommendations that take into account previously typed annotations by all users are made available on the "Enter your Text Note" field in the pop-up annotation window. Alternatively, users may choose from the list of annotations typed previously. A last collaborative annotation feature to be noted is the possibility of converting text annotations to captions in the local copy of the video.

3. FINAL REMARKS

The main features of the CWacTool have been presented from the "in use" point of view. The tool has been evaluated using several approaches: its use, its underlying architecture, and results from usability evaluations are detailed elsewhere [3].

Acknowledgments: We thank CAPES, CNPq, FAPESP, FAPEMIG, FINEP, CITIC/MCT.

4. REFERENCES

[1] Cattelan, R. G., Teixeira, C., Goularte, R., and Pimentel, M. G. 2008. Watch-and-comment as a paradigm toward ubiquitous interactive video editing. ACM TOMCCAP Vol 4, no. 4, article no. 28, pages 1-24.

[2] Motti, V. G., Fagá, R., Catellan, R. G., Pimentel, M. C., and Teixeira, C. A. 2009. Collaborative synchronous video annotation via the watch-and-comment paradigm. In EuroITV '09. ACM Press, pages 67-76.

[3] Pimentel, M; Motti, Teixeira, C., Faga Jr, R. Collaborative Synchronous Video Multimodal Annotation via the Watch-and-Comment Paradigm: evolution through evaluations. (under review)

The IITV Project, a video authoring tool based on recommendation.

Toni Bibiloni, Miquel Mascaró, Pere Palmer
Lab Multimedia - UIB
University of Balearic Islands
Crta Valldemossa Km 7,5
+34 971 173204
{toni.bibiloni, mascport@uib.es, pere.palmer@uib.es}@uib.es

ABSTRACT
The Interactive Internet TV Project is intended to be a web tool for publishing, indexing and viewing videos; that offers ontology based recommendations. The application has been developed and tested as a platform for e-learning in a university context and in Optiva Media Company. It combines expertise in education and television to develop a prototype, where the contents are generated using an automatic recommendation model based on personal profiles of each user; and that can be accessed from different devices: computers, mobile phones and so on.

Categories and Subject Descriptors
H.3.3 [Information Systems]: Information storage and retrieval. Information systems applications.

General Terms
Experimentation, Human Factors, Standardization.

Keywords
Content-based retrieval, MPEG-7, ontologies, recommendation, IPTV, Java, mobile devices.

1. INTRODUCTION
The current developments in video and digital television technologies, the Internet expansion and also the broadband technologies (xDSL and also iTV) have helped the sharing of interactive audiovisual contents. Thus, the number of digital libraries and the amount of audiovisual content available are growing and growing.

However, the high amount of audiovisual contents, and the format those contents have, represent a serious drawback in respect to the way the systems can handle them. Furthermore, the handling of the semantic data related to the audiovisual contents represents a new handicap, because of the complex relations that must be evaluated in order to offer the adequate information to each user; and thus it's very complicated offer this to the final user with a reasonable QoS. The problems are especially related to the audiovisual data recovery. The usual retrieval techniques are not suitable for practical use with digital video libraries.

The inefficiency and limitations of traditional retrieval systems has resulted in an increased demand for new techniques to manipulate the audiovisual data based on its contents. The content-based retrieval is considered one of the most complex challenges in multimedia environments, requiring a multidisciplinary research involving database engineering, computer vision, digital signal processing, artificial intelligence, pattern recognition, psychology and many other fields.

In this context our goal is to use the contents of the audiovisual data files to generate an automatic indexing based on its segmentation.

The video information probably offers the contents in the closer way, direct, accessible and easy to understand. Nowadays there exists an evident trend to use Television in non traditional form, for example Internet and mobile devices (e-Hand) are becoming a very suggestive alternative to the traditional TV devices, because of their connectivity and portability. Currently there are more potential users in the mobiles devices market than in the Internet; so we consider taking care of this distribution channel, and adapting our development to it, has a key role.

The Interactive Internet TV Project, in short the IITV Project, has as a main goal the creation of a broadcasting platform for digital audiovisual contents. The contents offered to the final user must fit some criteria; in order to do this the audiovisual data have to be semantically indexed and segmented. The segmentation has to be done in a semiautomatic way.

The proposed platform will be integrated by an interactive application to manage and distribute multimedia contents over internet TV, altogether with an indexing system capable to classify the digital content in a semiautomatic way and a mobile communications module allowing the digital broadcasting to mobile devices; thus the application multiple device oriented: e-Hand and Internet.

The application has been designed to be used in any context; in particular it can be useful to complement the E-learning platforms. The platform could also be used for governmental or social communication; or also for social networking or business and cultural promotion.

The educational environment generates a great volume of textual and audiovisual information. All of this information is semantically strongly related but at the same time it is physically decoupled. Thus all of this information has a very short life cycle, usually its own temporary duration, and without any reusability in the educational environment. The recovery, cataloguing and later

presentation of this content allows demonstrate the benefits of our proposal.

2. GENERAL ARCHITECTURE

The general architecture of the Interactive Internet TV Project is as shown in (Figure 1). The audiovisual information is stored in a DBMS. The information is stored as a raw data, in some video format. In order to have the semantic information available, a backoffice module allows adding, modifying or removing the semantic structure related (in our case study: professors, themes, activities, courses, and so on).

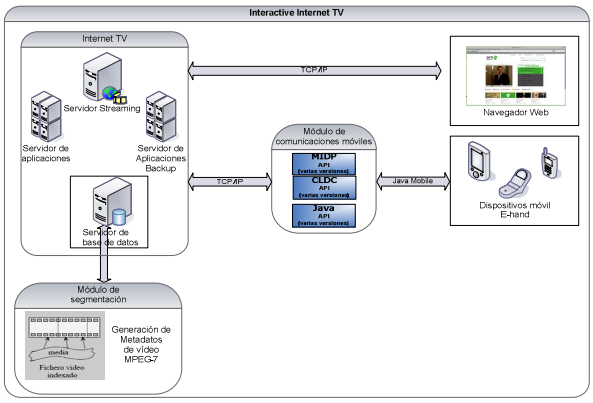


Figure 1. The platform general architecture.

With the segmentation module the audiovisual data can be related to the semantic structure. In this way every movie can be divided logically in scenes each one describing some kind of activity; where one or more actors (teachers, in our case study) perform their actuation (the activity). The scene also is related to some kind of semantic content (the theme or lesson given, or a part of it).

The information can be accessed via web, using a web browser, o via mobile devices, using the mobile communications module. This module is based on the JavaME architecture so the module can be used in the widest range of mobile devices.

Moreover there is another module, the recommendation module. This module is responsible for determining what content should be offered to each individual user. The content is gathered using specific search criteria for each user registered, that is: for each student. The information taken into account comes from the academic information. Thus, the contents offered to each user will be chosen according to their academic profile. Obviously, the search criteria entered by each user are also taken into account. Furthermore, the interaction with each user: what the user does and the value given to each media presented; is a feedback used in the recommendation algorithm.

3. THE SEGMENTATION MODULE

The aim of this module is to build a tool that can easily extract semantic information from audiovisual content. The audiovisual content information is easily interpreted by people, but it becomes extremely difficult to handle by computer systems, this is due both to the material, the content self, as to the way in which this information is represented, its format and size. When information is textual, only words, the scenario is relatively simple. But the audiovisual information has so many aspects that the number of nuances that may affect the interpretation of that information can be very big. So having an efficient and simple method to search, manage, acquire, process and retrieve information is critical.

The automatic information processing must be capable of extracting several features that allow generating descriptors that can easily identify each one of the components of the audiovisual information handled. There is now a standard, MPEG-7, which allows precisely this sort of description, see [1]. The different parts of MPEG-7 standard provides a comprehensive suite of tools for describing multimedia content that can create so-called descriptions that can be used to allow quality access to content, see also [2].

Therefore, the tool should be able to easily segment the audiovisual content supplied and also should provide a set of metadata expressed according to the MPEG-7; so each segment individual can be recovered with them. Accurately identify each segment requires that information on the content, and the information about the different elements generated by the segmentation system, will be stored in a Database Management System. So the information recovery can be done in an easy way. The (Figure 2) shows the aspect the tool has.

Figure 2. The visual segmentation tool.

After extracting the information, this must be stored in a database. The data and associated information can be retrieved efficiently, in response to issues like "*I want to see the proof of the theorem of Pythagoras*" or " *I want the French lesson given on the Thursday last week*". The search must take place among the descriptions to find those that match the question.

Another possible way of extracting information would by inference on requests already made, that is: once the user has made a request and has been responded, then the system automatically performs another search to identify similar information, for example other theorems of the same level/course; other subjects taught or lessons given on the same day as requested, or French lessons taught successive days. All this information could be offered to the applicant in order to facilitate their work.

4. THE VIEWING MODULE

The viewing module is a simple way to have access to the semantically segmented information. The user can select several

search criteria and the response is the set of matching segments, see (Figure 3). An ontology-based recommendation system is the responsible to gather the segments that fit better with the criteria and the user profile, see [3].

Figure 3. The viewing module.

The set of segments with the corresponding semantic data can be viewed sequentially, one after other from the first one to the last, or alternatively the user can decide step over some of them, go back, and so on. The order in which the segments are sorted depends on the search criteria entered by de user and also the ratings given previously to each segment, so the best rated are the chosen ones to be shown before. The ratings must be collected from the users, thus it is possible with the segment viewer to rate the degree of satisfaction of each segment viewed.

Figure 4. Viewing a video on an e-Hand device.

4.1 The mobile communications module

The viewing module has been designed to have two possible execution environments. The first one is a web application intended to be executed on a computer, the second one is a mobile application designed to be executed over a e-Hand device, for example a mobile phone.

The functionality is exactly the same in both cases; see (Figure 4) to view how a video is displayed on an e-Hand display. Maybe the main difference between both modules is the screen resolution and the video formats; these topics are a severe handicap for the mobile devices; because usually they don't have anything comparable to a standard computer neither in screen resolution nor video formats supported.

5. IMPLEMENTATION DETAILS

The different modules developed in our applications has been developed using the Java programming language, using different platforms depending on the module (Java EE, Java SE, and Java ME). The computer viewer and the visual segmentation tool use shockwave flash technology. The database has been implemented over MySQL. The MPEG-7 data generated by the application has been validated with the NIST service (http://m7itb.nist.gov/M7Validation.html).

6. CONCLUSIONS

In first place, our project will improve technologically the on-line training platforms, refining their formation skills and also improving the broadcasting processes and the multimedia content recovery. Also, our contribution can increase the services offered by the different digital libraries of the educational centres, in which our development will be installed. Moreover, the integration with the mobile communications module will expand the accessibility range, allowing the access to the contents broadcasting service to the e-Hand users.

And, as discussed in the introduction, the development can be applied in a wide range of environments, not only in the educational one.

7. ACKNOWLEDGMENTS

Our thanks to Optiva Media, our partner in this project; and also to our colleagues from the *Laboratorio de TV Digital Interactiva* at the University of Vigo, theirs is the recommendation module. This work has received funds from the Spanish *Ministerio de Industria, Turismo y Comercio*.

8. REFERENCES

[1] B. S. Manjunath, P. Salembier, T. Sikora. 2002. Introduction to MPEG-7: Multimedia Content Description Interface, S. Mullender, Ed. John Wiley & Sons. ISBN: 978-0-471-48678-7

[2] Martinez, J.M.; Koenen, R.; Pereira, F.; , "MPEG-7: the generic multimedia content description standard, part 1 ," Multimedia, IEEE , vol.9, no.2, pp.78-87, Apr-Jun 2002 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=998074&isnumber=21537

[3] Y. Blanco-Fernández, J. Pazos-Arias, A. Gil-Solla. 2009. Semantic Web Technologies in the Service of Personalization Tools. Handbook of research on social dimensions of semantic technologies and web services. IGI Publishing, 68-87

Controlling Interactive Digital Television Applications by Brain-Computer Interface

Taciana Saad Rached
Embedded Systems and Pervasive Computing Laboratory
Federal University of Campina Grande
Campina Grande, PB, Brazil
taciana.rached@ee.ufcg.edu.br

Danilo Freire de S. Santos
Signove Tecnologia S/A
Campina Grande, PB, Brazil
danilo.santos@signove.com

Angelo Perkusich
Embedded Systems and Pervasive Computing Laboratory
Federal University of Campina Grande
Campina Grande, PB, Brazil
perkusic@dee.ufcg.edu.br

Hyggo Almeida
Embedded Systems and Pervasive Computing Laboratory
Federal University of Campina Grande
Campina Grande, PB, Brazil
hyggo@dsc.ufcg.edu.br

Maurus Marques de Almeida Holanda
Department of Internal Medicine
Federal University of Paraíba
Campina Grande, PB, Brazil
maurusholanda@hotmail.com

ABSTRACT

The development of assistive technologies has a major impact in the life of people with severe neuromuscular diseases. These technologies are important to provide or extend functional abilities of people with special needs, thus, promoting independent life, social inclusion and improving quality of life. In the context of Digital Television one of the main challenges is the design of artifacts that allow disabled people to interact with interactive applications. In this paper we propose the use of BCI-aware systems as a new interaction method between users and digital television applications. Moreover, this work presents an architecture to this proposed system that could be used for controlling any digital television interactive application. In this paper it is also presented a study case using the Ginga-NCL middleware.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

General Terms

Brain-Computer Interface

Keywords

BCI-aware pervasive systems, BCI, smart devices, digital television, disabled people

1. INTRODUCTION

Brain-Computer Interface (BCI) has been the subject-matter of recent studies in Human Computer Interaction (HCI) [1] in the context of both clinical studies as well as consumer electronics to enable people with impairments to control interactive applications. In this context, the recent availability of low cost consumer electronics devices to be used for BCI applications such as the Mindset headset from Neurosky[1], is a major trend to allow low cost BCI to promote the communication of motor disabled people with the environment.

Nowadays, the advent of wireless technologies and the increasing number of powerful portable smart devices such as personal digital assistants (PDAs) and Internet tablets make possible the design of systems where mobility is now a reality. Also, it is important to point out that user intervention can be as minimal as possible so that simple tasks can be performed with very few or no human interventions at all.

Ibrahim and Johansson in [2] and Lobato et al. in [5] discussed the use of multi-modal systems for users interact with TV applications. These systems combine visual and speech recognition modes for interaction with the DTV. In the first work, the system was designed by adding speech interaction to an already existing TV program guide. In the latter, speech recognition and synthesis systems were implemented in a mobile device, using its own audio inputs and outputs. The problem with the use of speech recognition for TV control is that the use of voice command in noisy environments cannot work properly. Besides, people with loss speech and motor disabilities cannot interact with the television using this type of system. SonarPen [9] is an ultrasonic pointing device system to control the DTV. The SonarPen system was used to perform few function controls such as channel, volume and brightness knob. The user chooses a control knob by focusing and changes the channel, volume and brightness by a rotating gesture while pressing the knob. Ultrasonic pens do not provide alternatives for interaction with the television system for users with special needs. Lin and Chen in [4] presented a system built with the middleware MHP (Multimedia Home Platform) compatible with most DTV systems supporting DVB (digital video broad-

[1] http://www.neurosky.com/

casting) standard [7]. Viewers can remotely command their own set-top boxes to record DTV programs and to watch pre-recorded ones, respectively, using numerous mobile appliances attached in an IP network such as mobile phones and laptops, among others. This type of system does not provide alternatives for interaction with the television system for users with special needs. Also, users with visual impairments may find it difficult to use these devices since they usually lack hearing feedback.

In the context of Digital Television (DTV), Interactive DTV applications for the Brazilian Digital TV System are implemented using the middleware Ginga[2]. Ginga allows the use and development of procedural and declarative applications. This paper presents a new artifact to allow disabled people to interact with Ginga Interactive Brazilian DTV applications. Thus, it is introduced an architecture for DTV control based on brain activity using a smart device and pervasive services, that way providing mobility and social integration for the user. Besides, an interactive DTV application developed using the Ginga-NCL language and integrated in the introduced architecture is presented.

The rest of this paper is organized as follows: The architecture for the development of DTV applications controlled by BCI-aware systems is presented in Section 2. Section 3 shows the BCI-aware pervasive multimedia systems design. The application to validate the architecture introduced in this work is presented in Section 4. Finally, Section 5 discusses conclusions and future works.

2. SOFTWARE ARCHITECTURE FOR A BCI BASED MULTIMEDIA DEVICE INTERACTION

Having in mind that one of the major objectives in the context of this work is to define an architecture sufficiently general for the control of BCI equipped DTV devices, in this section we focused on the details of such architecture. The reader should notice that it is presented a high level discussion of the software architecture framework due to space limitations in the length of this paper.

As can be observed from Figure 1, the proposed architecture introduced in this work has two major modules, namely: the *BCI module* (BCIM), and the *Data Mapping and Context Module* (DMCM).

In the BCI module the main functionalities are brain signals acquisition and processing. Several methods can be used to acquire brain signals. Electroencephalography (EEG) is the most used method in non-invasive BCI applications because it is portable, inexpensive and can be used in different kinds of environments [8]. Observe that with the recent advances in BCI research area, some companies are making available BCI modules with integrated controllers that can perform most of the processing needed.

After the acquisition, the signal is digitized and processed. Signal processing is divided in two steps, namely, features extraction and translation algorithm. The first one is used to extract the desired features of the signals. In the latter step, extracted features are translated into command signals.

The Data Mapping and Context module is used to transform processed signals into commands for controlling DTV systems. This transformation occurs based in the user profile and context data provided by the Profile/Context module, together with the features of the DTV application that the user desires to control (Features Selection module). The generated information is used by the Command module to map the user intention into commands to control the DTV system.

[2]http://www.ginga.org.br

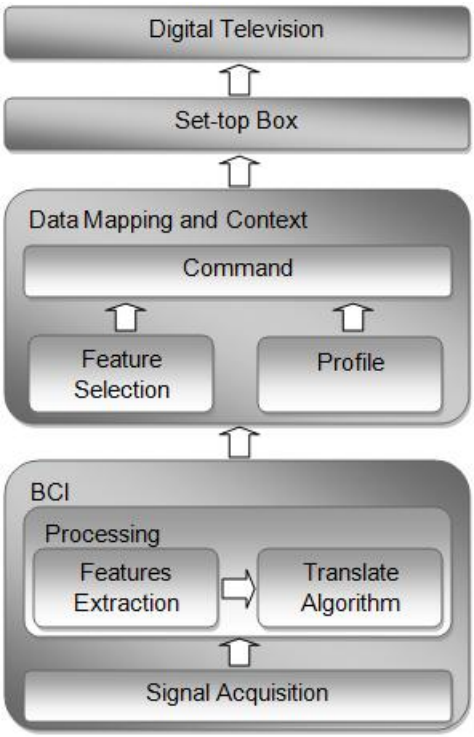


Figure 1: Architecture for BCI-Aware Pervasive Multimedia Applications

3. SYSTEM DESIGN

The system presented in this section was designed to allow people with severe motor impairments to interact with DTV applications. An illustration of the system including all of the components is shown in Figure 2.

EEG signals have several features in the frequency domain such as *alpha* waves (8 - 12.99Hz). *Alpha* waves increases when the person closes his eyes, and when he opens his eyes

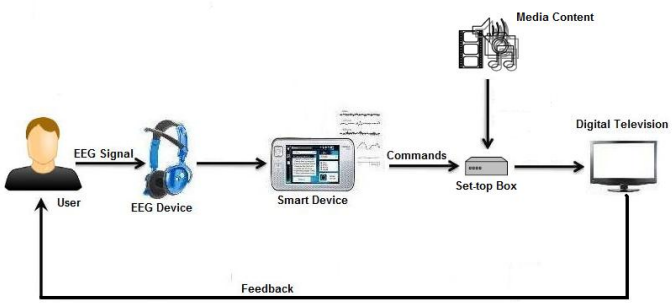


Figure 2: System Design for Multimedia Devices Control

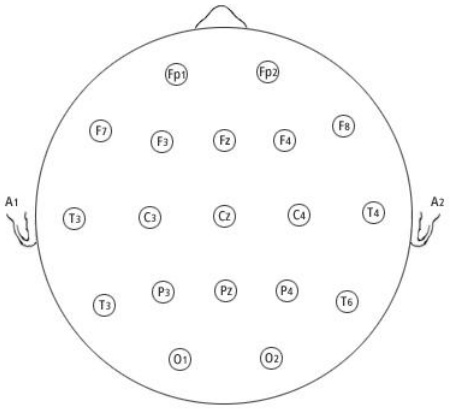


Figure 3: The International 10-20 system

these waves are decreased. *Alpha* waves generated from the brain occipital area during periods of closed eyes are the strongest EEG brain signals. They usually can be detected with the naked eye. Specially, *mu* waves (8 - 12Hz), a variant of *alpha* waves, can be found over the motor cortex (central scalp) that is reduced with movement, or the intention to move [10].

An EEG device measures non-invasively brain electrical activity of users by electrodes applied on the scalp according the International 10-20 system. The International 10-20 system is used to standardize the location of electrodes to easy comparison between subject studies [3]. This system is shown in Figure 3.

In our system design, the signal processing is realized by a embedded chip in the EEG device. A portable smart device is used as gateway between an EEG device and a DTV device such as a DTV set-top box. The portable smart device and the EEG device are connected through a Bluetooth interface. The portable smart device reads from the serial port (RFCOMM) the processed signals and translates their into commands to control interactive DTV applications. The use of a portable smart device is important mainly due to the limited DTV set-top box computing power. A network interface is used to connect the portable smart device and the DTV device. The DTV device uses a IP address to access the commands and maps these commands into remote control functions. Finally, an interactive DTV application runs in a set-top box to validate the system.

4. EXPERIMENTAL RESULTS

It was developed a BCI-aware pervasive system for controlling digital television application based on the architecture presented in Section 3. To validate the system, it was implemented an interactive application to demonstrate the interaction between the BCI-aware pervasive system and the DTV. In this application, the user can obtain information about the Soccer World Cup 2010 such as the last news, games, online surveys and buy tickets for the desired game using just his/her brain activity.

4.1 Hardware and Software Description

The brain activity was recorded in real time using a multichannel EEG equipment with three channels of Ag/AgCl electrodes. The electrodes are attached on the user scalp according to the International 10-20 system. We use electrodes in C3 and C4 sites to recognize the right and left arm movements respectively, and an electrode in the location O1 to detect whether the user closes his/her eyes or not. The reference (Ref) and the ground (Gnd) electrodes were mounted in the left ear.

A chip embedded in the EEG device was used to process brain signals. This chip was developed by NeuroSky [6] and is used to separate EEG signals in several components such as *alpha* waves. In our system implementation is used *alpha* components to determine the right and left movements, and closing or opening eyes.

An Internet Tablet (N810) was used to transform *alpha* waves into commands for controlling interactive DTV applications based on the user intentions. *Alpha* waves detected during closed eyes are transformed in a selection command and *alpha* components detected during the intention of right and left arm movements are transformed in right and left movement commands, respectively.

Instead of use a real set-top box device, it was used a Personal Computer (PC) which simulates a set-top box device running a Ginga-NCL middleware. The use of a PC device was made because the easy of development and controlling of software components in an open platform. Commercial set-top boxes uses proprietary systems, which make difficult software extensions and platform access.

As introduced before, it was developed a Soccer World Cup Ginga-NCL application to validate our system. This application was developed using the Composer software, which is a Ginga-NCL tool. The Soccer World Cup application make available information services about the Soccer World Cup for the user .

4.2 System Operation

In this system, users brainwaves are acquired by the EEG device. The system embedded in the EEG device extracts signals features and uses a translation algorithm to transform the EEG components in float numbers that represents their amplitudes. The EEG equipment communicates with the N810 Internet Tablet through the Bluetooth interface. Nokia N810 receives processed signals from the Bluetooth serial port (RFCOMM), and transforms those signals in three types of commands according to the user intention: left or right movement or selection commands. The DTV device connects with the Internet Tablet using an IP address to obtain the commands and to map these commands into remote control functions.

In the DTV device, during the transmission of a soccer game, an interactive application icon is shown in the screen to inform that there are interactive applications available for the user. The user can execute the application that has been presented in the television, in this case the Soccer World Cup application, by selection of the application icon. This selection is done when the user closes his eyes. During the closed eyes period, *alpha* waves appear strongest in the occipital brain lobe and then, when the user opens his eyes, these waves are reduced. This oscillation in *alpha* waves is measured by the electrode attached in the O1 location and processed by the chip embedded in the EEG device. Nokia N810 transforms *alpha* waves into a select command.

After the user selects the interactivity icon, the interactive application starts. The layout of the Soccer World Cup application is presented in Figure 4.

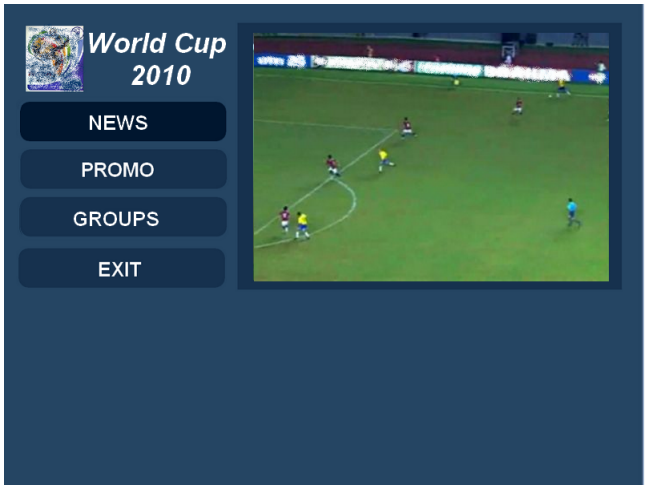


Figure 4: Screenshot of the World Cup 2010 Digital Television Application

When the user moves or imagines the arm movement, a variant of *alpha* waves called *mu* waves is reduced in the region of the central sulcus. When the movement is executed by the right arm, the reduction of the mu waves occurs in the brain left hemisphere. This reduction is recorded by the electrode attached in C4 location. These waves are transformed into a right movement command. In the scenario presented in Figure 4 the News item is highlighted, though when the user imagines the arm right movement the command to move for right is executed and the next item will be highlight. In the case of this application the next item is the Promo item.

When the user moves or imagines the movement of the left arm, occurs a reduction in the *mu* waves in the central sulcus located in the brain right hemisphere. The electrode applied in the C3 location is used to acquire this reduction. The Internet Tablet is used to transform these signals into a left movement command. In the scenario described above the Promo item is highlighted, but when the user imagines the left arm movement, *alpha* waves are translated into a left movement command. When the left movement command is executed the previous item will be highlight. In this case the previous item is the News item.

5. CONCLUSIONS AND FUTURE WORK

In this paper it was presented an architecture for the implementation of BCI-aware pervasive systems for controlling interactive DTV applications developed based on the Ginga standard. The main focus on the definition and implementation of the introduced architecture was to propose a new channel of communication between disabled users and DTV applications.

Also, we have discussed the development of simple applications providing users with more mobility, from the device point of view, and independence, allowing people with motor disabilities to control DTV applications anytime and anywhere. We developed a system to control a DTV application to validate the introduced architecture. Results of tests have shown that it is possible to insert people with severe neuromuscular diseases in the society with the development of this type of application.

For future works we plan to extend the introduced architecture to allow the interaction of BCI-aware pervasive systems with any multimedia device. In this sense, in order to explore novel approaches to be applied together with BCI, we plan to investigate the use of the user location information and user preferences to help in selection of the multimedia device that the user desires control.

6. REFERENCES

[1] T. Ebrahimi, J. Vesin, and G. Garcia. Brain-computer interface in multimedia communication. *IEEE Signal Processing Magazine*, 20(1):14–24, 2003.

[2] A. Ibrahim and P. Johansson. Multimodal dialogue systems for interactive TV applications. In *ICMI '02: Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*, page 117, Washington, DC, USA, 2002. IEEE Computer Society.

[3] R. Krepki, G. Curio, B. Blankertz, and K.-R. Müller. Berlin brain-computer interface-the HCI communication channel for discovery. *Int. J. Hum.-Comput. Stud.*, 65(5):460–477, 2007.

[4] C.-C. Lin and M.-S. Chen. On controlling digital TV set-top-box by mobile devices via IP network. In *ISM '05: Proceedings of the Seventh IEEE International Symposium on Multimedia*, pages 52–59, Washington, DC, USA, 2005. IEEE Computer Society.

[5] V. Lobato, G. Lopez, and V. M. Pelaez. MHP interactive applications: Combining visual and speech user interaction modes. Leuven, Belgium, 2009. EuroITV 2009.

[6] G. Rebolledo-Mendez, I. Dunwell, E. A. Martínez-Mirón, M. D. Vargas-Cerdán, S. Freitas, F. Liarokapis, and A. R. García-Gaona. Assessing neurosky's usability to detect attention levels in an assessment exercise. In *Proceedings of the 13th International Conference on Human-Computer Interaction. Part I*, pages 149–158, Berlin, Heidelberg, 2009. Springer-Verlag.

[7] U. Reimers. The DVB project-digital television for europe. In *DVB (Digital Video Broadcasting): The Future for Television Broadcasting?), IEE Colloquium on (Digest No.1995/142)*, pages 1/1–1/7, Jun 1995.

[8] A. Schlogl and C. Brunner. Biosig: A free and open source software library for BCI research. *Computer*, 41(10):44–50, Oct. 2008.

[9] M. Sohn and G. Lee. Sonarpen: an ultrasonic pointing device for an interactive TV. *Consumer Electronics, IEEE Transactions on*, 50(2):413–419, May 2004.

[10] H.-G. Yeom and K.-B. Sim. ERS and ERD analysis during the imaginary movement of arms. In *Control, Automation and Systems, 2008. ICCAS 2008. International Conference on*, pages 2476–2480, Oct. 2008.

Text Input in Digital Television: a Component Prototype

Diogo Pedrosa, Didier A. Vega, Maria da Graça C. Pimentel, Renata Pontin M. Fortes
SCC/ICMC/USP - Caixa Postal 668, 13560-970, São Carlos, Brazil
{diogo,davo,mgp,renata}@icmc.usp.br

ABSTRACT
The interactive Digital Television should prompt the development of many applications that require users to input text. Existing methods for text input are limited in terms of efficiency and user satisfaction. This work introduces both the design and the prototype implementation of a multimodal mechanism of text input in DTV in which three modes are offered: speech, mobile phone style and virtual keyboard modes. Interaction is performed with a remote control with a built-in microphone and a push-to-talk key. In the current prototype version, an adapted computer keyboard is used and the speech recognition engine is not implemented.

Categories and Subject Descriptors
B.4.2 [**Input/Output and Data Communications**]: Input/Output Devices---voice; H.5.2 [**Information Interfaces and Presentation (e.g., HCI)**]: User Interfaces – Input devices and strategies (e.g., mouse, touch screen), Voice I/O

General Terms
Design, Experimentation, Human Factors.

Keywords
Text Input, Digital Television, Interactive Television, Virtual Keyboard, Multi-tap, Predictive text.

1. INTRODUCTION
The emergence and evolution of Digital TV (DTV) standards have enabled the development of more complex applications. One of the limiting factors for application development is the difficulty of providing good usability based only on traditional remote controls, which have usually the directional arrows and the Ok, Back, colored and numeric buttons.

Good usability in interactive TV is particularly difficult to achieve in applications that require text input. Research challenges include support to multimodality [1, 4, 6, 9] and cultural differences [7, 8], and efforts have led to development of new devices [4] and of new usability evaluation methodologies.

In this paper, we present the design of a multimodal mechanism for text input in the DTV, along with a corresponding prototype and its usability evaluation (Section 2). We present the future works in Section 3 and our final remarks in Section 4.

2. DESIGN & PROTOTYPE
To design a mechanism for text input in DTV applications, we have first carried out requirements elicitation with users, followed by a literature survey in constraints of DTV in general. The project was conducted using techniques that aim to engage the end user in all stages of the design, known as User Centered Design (UCD) (see, for instance, [2, 5]).

2.1 Requirements Elicitation
Initially, a study to better understand the future users of the mechanism to be developed was conducted. At this stage, potential users filled questionnaires and experts in DTV and IHC were interviewed. The questionnaires were administered on paper (32 responses), in order to reach an diverse audience, and using a on-line surveys system (121 responses), which helped obtain a greater number of responses.

One of the questions asked what would be the best way to write a message to a friend using the television. The answers to this open question were very diverse and may be classified into: QWERTY keyboard, T9, speech, virtual keyboard using a simple remote control, virtual keyboard using a touch screen TV, virtual keyboard using a touch screen remote control, thought, pre-formulated phrases, writing on a paper whose text is recognized by the television, writing with a pen directly on television screen, and using a mobile phone connected to the TV.

A concern regarding the need to offer more than one alternative to text input could be noticed, as in the examples (our translation):

- "*Using speech when I'm alone and using T9 when I'm in an ambient with other people.*"
- "*Speech would be ideal, but I think we could correct some possible errors or [make some] modifications using a keyboard, for example*"
- "*I would like to be able to choose: i) If I'm in the living room with my mother in law: conventional keyboard; ii) if I'm with other people: speech to text.*"
- "*Using speech and phonemes recognition or, to a reality closer to ours, a remote control with a LCD touch screen...*"
- "*... but there must be other ways, which are accessible from those who are dumb or are without voice, for example.*"

Another recurrent concern was regarding to the interference in the environment and TV sound, in cases in which the text input is performed using speech, as in the following examples (our translation):

- *"... I know there are limitations and difficulties, such as external interference from other people, background noise, etc."*
- *"... but there may be noise problems…"*

- *"... without interfering with the program audio that the person is watching."*

This step was crucial to help us to understand the importance of offering alternative ways of entering text, in order to meet the needs of different user profiles and to be flexible enough to be consider different environments.

Still aiming at better understanding the problem, we conducted a survey on the ways that text input is currently used in TV systems, and a review of recent scientific papers. The literature shows that the problem involves complex issues, such as multimodality to meet a wide range of users [1, 4, 6, 9], cultural differences [7, 8], development of new physical devices [4], and the use of a usability evaluation methodology specific to the area [3].

The two steps allowed us to define the functional requirements and key usability criteria to be considered. Nine functional requirements for the system were elicited, addressing basic issues related to inserting and deleting characters, and the substitution of letters both using speech and remote control. The usability criteria considered the most important in the context of the project were: (1) *Familiarity*: users should encounter familiar writing concepts from the first interactions with the proposed mechanism, (2) *Substitutivity*: complementary forms of text input should be provided, and (3) *Responsiveness*: even though DTV decoders usually have little processing power when compared to computers, the system should be fast enough to allow users to notice changes during input.

2.2 The prototype in use

Two interfaces were independently designed, and their strengths combined to create a third interface. Next, a functional prototype was implemented in Java, aiming a short development time. The component has three modes of text input: speech mode, mobile phone style mode and virtual keyboard mode. Figure 1 shows the four major states of the text input component prototype.

Figure 1. Four main states of the text input component prototype (in Portuguese): Speech mode in Dictate state (1), mobile phone style mode (2), virtual keyboard mode (3) and virtual keyboard mode in cursor movement state (4)

The component has three text input modes: speech mode, mobile phone style mode and virtual keyboard mode. The speech mode would allow the user to write in the selected text box dictating words (no speech recognition was implemented). This functionality was tested using the Wizard of Oz technique. In addition to the Dictate state, shown as *Ditar* in Figure 1 (1), the speech mode of the component has also a state where only editing commands are recognized by the speech recognition engine (not available in the current version).

The mobile phone style mode (Figure 1 (2)) allows users to write in the selected text box using the numeric keys of the remote, just as it is traditionally done in telephones. It should allow both multi-tap and predictive text (only the first mode is currently available).

Finally, the virtual keyboard mode (Figure 1 (3)) allows writing text by selecting the desired characters using the arrows plus the OK keys. By selecting the button on the bottom right of the keyboard, the function of the arrow keys on the remote control changes to let the user to move the cursor in the text box and, therefore, the buttons on the virtual keyboard become disabled, as shown in Figure 1 (4).

Interaction with the real system should be performed with a special remote control containing a built-in microphone and a push-to-talk key that must be pressed to activate the speech recognition engine. In user tests with our prototype, the input data was performed using an adapted computer keyboard where some keys were relabeled and unused keys were covered with adhesive paper, in order to not confuse the user (Figure 2).

Figure 2. Adapted keyboard to interact with the prototype

2.3 Usability evaluation of the prototype

So far, the latest stage in our project was the usability evaluation of the prototype. Three techniques were used: Heuristic Evaluation, Cognitive Path and user testing using the techniques of thinking aloud (think aloud) and Wizard of Oz.

The tests were performed by 4 pairs of users, also following one of the recommendations of Flores et al. (2008), who note that tests pairs allows individuals to express more naturally their actions and opinions. The Heuristic Evaluations were performed using the general heuristics proposed by Nielsen and Mölich[1].

[1] The original list of 9 heuristics from Nielsen and Mölich (1990) has been refined by Nielsen (http://www.useit.com/papers/heuristic/heuristic_list.html)

The evaluations have identified a number of problems in the proposed mechanism. Seven of them, due to greater severity, were used to create a list of change recommendations. The main problems were:

1. The speech mode has a different activation mechanism from the others, but this is not indicated in the interface. There was not a consensus among the evaluators if this mode should or not be activated only by clicking on the push-to-talk button, as it is now. However, if this happens, that difference must be clearly indicated in the interface.
2. The activation mechanism of the mobile phone style mode and the virtual keyboard mode should be better crafted. The buttons that allow the activation are distant from the tabs that indicate which mode is active. The most intuitive seems to be to use their own tabs as the activation mechanism.
3. The mobile phone style mode does not give a clear indication that the buttons shown are merely illustrative. The buttons of the interface are shown only to serve as a guide for users who use remote controls with no letters printed on keys. This causes a large amount of error situations, in which users try to use the arrow keys to focus one of the illustrative buttons.
4. The term "Celular" (cell phone), used to indicate the text input style often used in mobile devices, may not be easily understood by users, who may think it refers to their own mobile phone or to a cell phone call.
5. The "BACK" button caused a lot of frustration because it could almost never be used. This may have occurred because the component has been evaluated outside the context of an application, where it would certainly have clearer functionalities. Another problem caused by this key was the association made between the key and the backspace key of a traditional computer keyboard.
6. The key used to enter and to exit the cursor movement state, accessible from the keyboard mode, was not the same. The key to enter was the "OK" key, but the key to exit was the "BACK" key. This was extremely counterintuitive. The possibility of inserting new lines using the "OK" key, offered when this state is activated, is not worth the amount of errors generated.
7. The interface of the speech mode is not clear enough and causes a lot of problems. The use of two columns of commands in the Dictation state makes the user associate the right column with the Commands state. The interface is not clear on how to access the Commands state, nor indicate clearly the difference between the Dictate state and Commands state. Not even the list of available commands in the Dictate state is satisfactory.

In addition to the problems identified in the design of the text input component, some serious problems on the specific prototype implementation ended up gaining more prominence in the evaluations than it was expected. They will be listed here only to its impacts be taken into account during the analysis of test results:

- The cursor that indicates where the next character will be inserted is not shown in any of the input text modes.
- The adjustment made on the numeric keypad resulted in the "Num Lock" key being used to insert symbols. This made the rest of the numeric keypad stop working whenever the symbols key was pressed an odd number of times.
- The predictive text functionality (T9), despite not being implemented, is indicated by a label on the interface.
- The tests of the speech mode, using the technique of the Wizard of Oz, did not allow the text to be inserted into the text area of the prototype. A parallel screen was used.

3. FUTURE WORK

In order to solve the main problems of the designed component listed in the previous section, the following changes should be made for the next version:

- Allow the three text input modes to be activated by selecting the corresponding tab. Thus, the “Celular” (“Mobile Phone”) button is removed from the virtual keyboard mode and the “Teclado” ("Keyboard") button is removed from the mobile phone style mode. When the virtual keyboard mode is activated, the focus can move freely between the character buttons and the two other tabs. When the speech or mobile phone style mode are activated, the focus can move only between the two other tabs. The button push-to-talk should be preserved and the user should be able to press it regardless of the active mode, which makes it a shortcut to the speech mode. However, if it is pressed while the speech mode is not active, when released, the previous mode should be activated again. In addition, to increase the clarity and to consider users who do not speak English, it should be relabeled to "Segure para falar” (“Hold to talk”). These changes aim to solve the problems 1 and 2, and also have some impact on problem 3, as it does not let the focus fixed on a single button while the mobile phone mode is active.
- The graphical interface of the mobile phone style mode should be improved so that no doubt remains that the buttons shown are merely illustrative. This change also aims to solve the problem 3.
- The tabs that identify the text input modes should be bigger to allow each mode be identified by more than one word. The mobile phone mode would be called “Estilo celular” (“Mobile phone style”), the virtual keyboard mode would be called “Teclado virtual” (“Virtual keyboard”). The name of the speech mode could remains the same. This aims to eliminate possible confusion explained by problem 4.
- The "BACK" key should be relabeled to "VOLTAR" (“back”) to consider users who do not speak English

and also solve the problem 5. If time is available, the context of the application should be used in the prototype, to make it clear what is the function assigned to the BACK key.

- The OK button should also be used to exit the cursor movement state. This would solve problem 6.
- The organization of the items shown in the speech mode should be improved, to make it clear to the user that "Dictate" is just one of the possible states of this mode. The command "Comandos" ("Commands") should be listed along with other possible commands of the Dictate state. Also, the command "Ditar" ("Dictate") should be listed along with other possible commands of the Commands state. The "Apagar linha" ("delete line") command should be added to the list of commands of the Dictate state. The commands that require some extra word, as is the case of "Insert symbol" and "Insert number" should indicate that in the interface. All these changes address the problems related to the speech mode, and grouped in item 7.

In future work, we intend to implement a new prototype that includes the suggested changes and targeting a set-top box (STB), so that new evaluations may be carried out, taking into account previously ignored factors, such as performance and use of a real remote control.

4. FINAL REMARKS

The prototype has undergone usability evaluations by means of Heuristic Evaluation, Cognitive Walkthrough, and user testing via Think Aloud and Wizard of Oz techniques. Although several problems have been identified, the elicited requirements were met in the design of the component. In the prototype, the speech mode and the predictive text functionality of the mobile phone style mode were not implemented due to time restrictions.

The component also met two usability criteria: *familiarity* – the deployment of traditional text input methods from other contexts, such as mobile phones, touch screen devices and video games, simplifies the first contact of new users – and *substitutivity* – since the designed component allows users to provide text for DTV applications in 3 different ways. The responsiveness criteria could not be verified because it requires a more mature prototype. Future work will tackle the limitations from the current version.

5. ACKNOWLEDGMENTS

We thank the following organizations for their support to our work: CNPq, Fapesp, Finep, Capes, CTIC/MCT. We also thank the subjects who participated in the evaluation studies.

6. REFERENCES

[1] Cox, A. L., Cairns, P. A., Walton, A., and Lee, S. 2008. Tlk or txt? Using voice input for SMS composition. Personal and Ubiquitous Computing, Springer London, 12, 567-588.

[2] Dix, A., Finley, J., Abowd, G., and Beale, R. 2004. Human-computer interaction. Person Education Ltd, 3rd edition, 318p.

[3] Flores, L. V., Faust, R., and Pimenta, M. S. 2008. Definindo uma proposta para avaliações de usabilidade de aplicações para o sistema brasileiro de TV digital. IHC '08: Proceedings of the VIII Brazilian Symposium on Human Factors in Computing Systems, Sociedade Brasileira de Computação, 88-97.

[4] Nakatoh, Y., Kuwano, H., Kanamori, T., and Hoshimi, M. 2007. Speech recognition interface system for digital TV control. Acoustical Science and Technology, 28, 165-171.

[5] Norman, D. a. The Psychology of Everyday Things. Basic Books, New York, 1988. In paperback as The Design of Everyday Things. Doubleday, New York, 1990.

[6] Patel, N., Agarwal, S., Rajput, N., Nanavati, A., Dave, P., and Parikh, T. S. 2009. A comparative study of speech and dialed input voice interfaces in rural India. In Proceedings of the 27th International Conference on Human Factors in Computing Systems, ACM, 51-54.

[7] Soronen, H., Turunen, M., and Hakulinen, J. 2008. Voice Commands in Home Environment - a Consumer Survey. In Proceedings of Interspeech 2008, pp. 2078–2081.

[8] Tan, G., Brave, S., Nass, C., and Takechi, M. 2003. Effects of voice vs. remote on U.S. and Japanese user satisfaction with interactive HDTV systems. Proc. Proc. CHI '03 extended abstracts on Human factors in computing systems, ACM, 714-715.

[9] Wittenburg, K., Lanning, T., Schwenke, D., Shubin, H., and Vetro, A. 2006. The prospects for unrestricted speech input for TV content search. AVI '06: Proceedings of the working conference on Advanced visual interfaces, ACM, 352-359.

Demonstration of the GAMECAST Interaction Environment

Thomas Schmieder
University of Applied Sciences
Mittweida
Rochlitzer Straße 50
09648 Mittweida, Germany
004917662399614
Thomas@gamecast-tv.com

Benny Baumann
University of Applied Sciences
Mittweida
benbe@geshi.org

Christoph Schubert
University of Applied Sciences
Mittweida
Cschube5@hs-mittweida.de

ABSTRACT
This paper presents the interaction environments of the Gamecast system. Gamecast uses virtual worlds (just like in multiplayer online games) to integrate players through their avatars as actors in the plot of a computer-animated TV series. The entire plot is filmed in the virtual world enabling scripted scenes and player content to seamlessly blend both visually and in terms of content. The players are in a position to influence the development of the plot through their interaction. The virtual world thereby becomes a virtual stage on which a cross-medial format emerges.

Keywords
converged media, computer games, future media, GAMECAST

1. INTRODUCTION
The production of the Gamecast series does not take place live during broadcast. Instead, in a specific time frame prior to the broadcasting of an episode, scenes from the player interaction are recorded as raw data and post-processed by a production team. The scenes are enriched with specially created elements which involves re-dubbing, introducing facial expressions and gestures, and the re-rendering of scenes in a better graphical quality. These are then integrated into the series at pre-determined positions.

In order that game scenes can be integrated into the plot of an episode, "gaps" are left in the storyline at certain points in the plot strand, which can be filled with post-produced scenes which have developed out of the player interaction. A further possibility is to "provoke" social developments in the game world. By undergoing plot related missions, or reacting to the provoked events, the players create a wide range of potential scenes which can then be selected from. To this end, Gamecast provides an analysis system which divides the player interaction into scenes and which can identify "suitable" scenes. So that this "video game film" is nevertheless interactive, the developments in the game world (such as, for example, a social development or the dominance of a group of players) can be analysed by the production team, taken up and integrated in the plot. For this reason also scenes which have been pre-produced can be introduced in a flexible way and adapted as necessary.

2. GAMECAST INTERACTION SPACE
The interaction concept of Gamecast is based on playful access to interaction and acting. This principle is called competitive acting [1]. The aim is to create structures and virtual interaction rooms in which lay persons can interact with each other by way of their avatars and at the same time appear as authentic virtual performers. In order to achieve this, the system must realise competitive situations (quests, missions) on the virtual stage. In these situations the players can concentrate on playing instead of acting out a particular role [1]. Players can invent for themselves an individual character in the game-based interaction environment, find their own role in this world or join one of the acting parties [2]. In both free interaction and in plot-related quests they can continue an episode and thereby influence the progress of the plot. The players can control their avatars from the first-person perspective by using a mouse, keyboard, or optionally also a gamepad. To avoid players having a sense of being limited but at the same time ensuring that the flow of interaction remains predictable, specific game targets must be set in the scenes, which the player can reach or not reach [1]. The players can move freely, choose tactics and thereby reach their goal. This competition does not necessarily have to result from a fight; it can be construed differently according to the situation. This model can therefore also apply to a sporting or social scene.

3. DEMONSTRATION
The starting point is two opposing parties, on the one side the police as a representative of the surveillance state, and on the other side the group of revolutionaries. From the extensive story a scenario in a desert landscape is selected to serve as an exemplary demonstration of the game mechanism. Besides the characteristics of the game parties there are a number of routes which can be taken towards the mission's goal. The choice of routes influences the actions of the other players, but a group is not limited to one route; it can use several. The revolutionaries'

goal is to obtain confidential information about a mental experiment. The task of the law enforcement forces is to prevent the intruders from doing so. Each of the game parties has different methods and routes available to them. The police use non-lethal weapons. Dependent on the approach taken by the players, the police can barricade themselves in surrounding the information, split up strategically, or position themselves in small groups on important defence lines. The revolutionaries only possess homemade weapons such as batons, grappling hooks and throwable objects. In order to avoid the defence installations cameras can be destroyed or manipulated. Furthermore, besides the direct route through the front door, further routes are available. This openness in approach leads to the players being encouraged to try out different courses of action and in this way different interactive scenes are produced. The assessment of the protagonists' actions affects the individual progression of a player as well as the overall scenario. In this way the social structure of the game world is promoted and the player is motivated to think about his behaviour as his actions have consequences.

4. PLAYER-OBJECT INTERACTION

In the same way that an action carried out between players becomes a specific dramaturgically meaningful game move [1], dramatic importance must also be attached to the interaction with objects. Just as the player becomes a virtual performer, significant objects become props for the purpose of the plot. So as to increase the dramatics also a notable emphasis is placed on the interaction with objects. This means that players must utilise these prudently and pay particular attention to their use. The plot itself has to become part of the game experience. To this end objects are linked to appropriate reactions (events, triggers) which, for example, enable a mini-game when a security system is triggered, and with which this can be either deactivated or activated. Objects contain pre-defined functionality which is enabled according to the environment and the plot. In this way the deploying of an object can trigger dramaturgical changes in the game situation and alter the win/lose conditions for all players. The use or disclosure of an object can be linked to a radical change in the course of the game, for example when a mobile phone is used in a confidential meeting to extract information from the opposing side. Objects and actions linked to them thereby gain dramaturgical significance: whether or not they are used sends a message to other users. Quests are set up in such a way that they promote a distribution of tasks amongst the players, each role representing a separate stage in the game. Despite divided roles only coordinated actions will lead to success. Violence leads to an advantage in the short term, on a global level it leads to disadvantages and a new game situation. To achieve this, variables are calculated by the game logic, which decide on not only the outcome of the mission, but also the difficulty level. Particular behaviour on the part of the players changes the variables. In scripting the NPCs attempts are made to mirror realistic behaviour and rules. It is decisive that in the case of intentionally disruptive behaviour by the players warnings are issued so as to steer the player back to the plot within the game world. NPC interaction will serve as a regulating element here.

Figure 1: Screenshots of the "Research Institution" Level

5. REFERENCES

[1] Schmieder, T., Wierzbicki, R. J. (2009) Competitive Acting - Issues on Action, Interaction and Acting in Converged Media, In Proceedings of the Academic MindTrek Conference 2009, Tampere, Finland, 30 - 2 Oct 2009, ACM Press.

[2] Schmieder, T., Wierzbicki, R. J. & Lugmayr, A. R. (2008) GAMECAST®: A Cross-Media Game And Entertainment System, In: Lugmayr, A. et al. (eds.), TICSP Adjunct Proceedings of EuroITV 2008, Changing Television Environments, Salzburg, Austria, 3-4 July 2008 (pp. 157-161). Retrieved March 30, 2010, from http://www.fgam.de/papers/2008/GAMECAST(R)_A_Cross-Media_Game_and_Entertainment_System_(2008).pdf

SUMA & T-Maestro: Integrating personalized t-learning through standards

Silvia Bastos
Laboratorio de TVDI, Universidad de Vigo
Campus Universitario s/n
36300 Vigo, Spain
+34 986 813901
sbastos@det.uvigo.es

Rebeca P. Diaz
Laboratorio de TVDI, Universidad de Vigo
Campus Universitario s/n
36300 Vigo, Spain
+34 986 813469
rebeca@det.uvigo.es

Ana Fernández
Laboratorio de TVDI, Universidad de Vigo
Campus Universitario s/n
36300 Vigo, Spain
+34 986 813868
avilas@det.uvigo.es

Gonzalo Jiménez
Laboratorio de TVDI, Universidad de Vigo
Campus Universitario s/n
36300 Vigo, Spain
+34 986 813901
gonzalo@det.uvigo.es

ABSTRACT

The SUMA[1] project is part of the INES technological platform eLearning working group and has been subsidized by the Ministry for Tourism, Industry and Trade, as part of PROFIT-the Avanza Plan's Industrial ICT Policy.

This project has developed a system for integrating eLearning services based on standards, which allows the incorporation of tools and services into a platform, in order to respond to the training needs of both the business and academic spheres, and to facilitate the integration of existing structures into the many fields that require the implementation of eLearning solutions. The basic aim of the project is creating a services layer that enables the separation of the eLearning base platform (Moodle or Sakai)[2] from the new applications and functions to be incorporated.

One of these new features aims the multimodal access to the eLearning platform through different devices (IDTV, Web, mobile…). IDTV's laboratory at the University of Vigo has focused in these services, providing multimodal access and designing contents adapted to user's profile and reception device. The demo presented in this paper shows the SUMA services for multimodal access trough Web and IDTV.

Categories and Subject Descriptors

K.3.1 **[Computer and Education]**: Computer Uses in Education – *distance learning, collaborative learning.*

H.5.1 **[Information Interfaces and Presentation (I.7)]**: Multimedia Information Systems – Video (e.g., tape, disk, DVI)

General Terms

Design, Experimentation, Human Factors, Standardization.

Keywords

T-learning, adaptation, personalization, IDTV, learning systems.

1. INTRODUCTION

The arrival of Interactive Digital TV makes the access to distance education easier, practically all European homes have at least one television set, whereas the penetration of Internet enabled computers is lower than TV. Apart from wide-world usage, TV is considered by the viewer trustworthy in reference to broadcast content and easy to operate. These conditions are an ideal starting point for TV-based interactive learning, referred to as t-learning.

Through personalization, it is possible to attract users towards education, choosing the most appropriate contents for them and adapting learning contents to their preferences and learning background. Through the combination of education and entertainment, it is possible to approach education to viewers.

2. SYSTEM DESCRIPTION

The Open Knowledge Initiative (OKI)[3] develops specifications that describe how the components of a software environment communicate with each other, OKI specifications enable sustainable interoperability and integration by defining standards for Service Oriented Architecture. The OKI project has developed the Open Service Interface Definitions (OSIDs), that defines important components of a SOA as they provide general software contracts between service consumers and service providers. This enables applications to be constructed independently of any particular service environment, and eases integration.

SUMA consists in three levels: the base LMS (Moodle or Sakai), SUMA OKI bus and applications that access to OSID services (see Figure 1). SUMA OKI bus offers an eLearning integration service layer of the OKI specification that provides a common standard interface for connecting applications regardless of the underlying LMS. This enables application integration paradigm using APIs that allows integrating implementations for different programming languages and adding different protocols. Our applications have been built over this integration layer.

One of the SUMA project aims is to provide multimodal access to adaptable content, to achieve this, the IDTV's laboratory have developed three fundamental entities:

[1] http://www.ines.org.es/suma

[2] http://moodle.org and http://sakaiproject.org

[3] http://www.okiproject.org

- ReloadTotal: the authoring tool. The courses created are defined in XML language that allows them to be independent of underlying technology.
- T-Maestro Intelligent Tutoring System: responsible for the component adaptation to the user profile.
- T-Maestro Interpreter: responsible for the adaptation to the reception device.

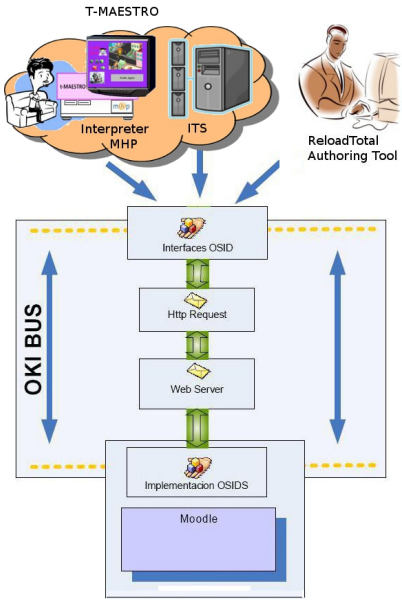


Figure 1: SUMA Architecture

2.1 ReloadTotal

The ReloadTotal is a tool to create adaptive learning elements based on SCORM (Sharable Content Object Reference Model), this tool adds new functionality to RELOAD Editor[4], that allows specifying different organizations for SCORM course. This authoring tool includes the submodule SCOCreator Tool which allows creating self-adaptive SCOs (Sharable Content Object).

With this tool we hide technical details to pedagogy experts and allow defining adaptive courses based on multimedia and using a media-driven strategy.

2.2 T-Maestro

The platform T-Maestro adapts at reception the eLearning contents to the user's preferences and reception device by combining the ITS and the interpreter. The interpreter is the responsible for the proper display according the access device and access to the eLearning platform to obtain the adaptable content. The ITS, performs the adaptation of the course according to user's profile. Thus, T-Maestro shows the user a course in which the content is adapted to his profile and his reception device.

3. DEMO SCENARIO

In the demonstration scenario (Figure 2) two roles are distinguished: course creators and students.

[4] http://www.reload.ac.uk/

Course creators are in charge of designing adaptive t-learning elements (pedagogical units or whole courses) by using the authoring tool we have developed, ReloadTotal.

The student has to ways to interact with the course, he can access to the adaptive courses through a web interface or an IDTV interface.

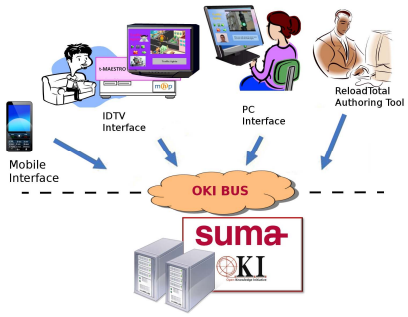


Figure 2: Demo Scenarios

A demo video is available at: http://idtv.det.uvigo.es/proyectos/t-learning/en/index.html

4. CONCLUSIONS

We have developed a pilot course about cooking, we have chosen because it is an appropriate subject for this type of training: very close to entertainment. They have participated a group of 40 heterogeneous users (both genres, different ages and educational levels), and we have obtained some conclusions by observing their reactions and by some questions results.

Overall, the study has identified that there is a big potential for utilizing the interactive digital TV solution for increasing learning opportunities in the home, particularly as an alternative solution to utilizing an Internet-enabled computer, 95% of respondents said they would use this application in their homes. This is interesting to obtain a broad participation in sectors of the population not familiar with the PC environment because 85% of respondents said they found easy to use the remote control, this is an important aspect for overcoming the digital divide.

A 60% of respondents said they would use this service on weekends, at night or in their leisure time, therefore, it should be used the playful nature of TV to make learning more attractive and to engage TV viewers in learning.

As a future line of work we are going to guide further studies on adaptable t-learning for elderly and we are developing the access to the adaptive courses trough mobile devices.

5. REFERENCES

[1] Rey-López, M., Díaz-Redondo, R. P., Fernández-Vilas, A., Pazos-Arias, J. J., López-Nores, M., García-Duque, J., Gil-Solla, A., and Ramos-Cabrer, M. 2008. T-MAESTRO and its authoring tool: using adaptation to integrate entertainment into personalized t-learning. *Multimedia Tools Appl.* 40, 3 (Dec. 2008), 409-451. DOI= http://dx.doi.org/10.1007/s11042-008-0213-4

Users in the loop: browsing archival video with advanced video retrieval technology

Johan Oomen
Netherlands Institute for
Sound and Vision
Sumatralaan 45, 1217 GP
Hilversum, The Netherlands
+31 624151991
joomen@beeldengeluid.nl

Roeland Ordelman
Human Media Interaction
University of Twente
Box 217, 7500 AE
Enschede, The Netherlands
+31 534892954
rordelman@beeldengeluid.nl

Cees G.M. Snoek
ISLA, Informatics Institute
University of Amsterdam
Science Park 107, 1098 XG
Amsterdam, The Netherlands
+31 205257528
cgmsnoek@uva.nl

Bauke Freiburg
Video Dock
Panamalaan 1b, 1019 AS
Amsterdam, The Netherlands
+31 202620838
bauke@videodock.com

ABSTRACT
The Dutch Pinkpop festival, the oldest annual rock festival in the world, celebrated its 40th birthday in 2009, giving rise to a number of festivities. One of these was the launch of the website www.hollandsglorieoppinkpop.nl ('Pride of Holland at Pinkpop'), an experimental video search application offering an innovative way to browse through a collection spanning 40 years of concert videos dug up from the archives of the Netherlands Institute of Sound and Vision.

1. SEMANTIC VIDEO CONCEPT DETECTION
The application is a unique publicly accessible demonstration of the application of state-of-the-art multimedia retrieval technology. It allows visitors to browse concert recordings and interviews based on automatically detected visual concepts and speech transcripts respectively. See Figure 1. To accumulate feedback on the performance of the technology, the user is challenged to flag both hits and misses. One out of ten visitors of the thousands that visited the website the past month really did so, providing valuable information for further improvement of the technology.

The technology has been developed over the past years by Dutch Universities, notably within the scope of the MultimediaN research programme. The main component is the MediaMill semantic video concept detection engine developed by the University of Amsterdam. The engine translates pixels to text, enabling the automatic labeling of visual concepts in video such as ‘guitar player’ and ‘drummer’.

The concept detection was trained to detect visual concepts using manually labeled examples. Thousands of features, related to the form, color and texture of images are processed in order to automatically create a link between a given segment and the concepts in the semantic video search engine. As a result, users can quickly browse to various concepts in a concert video using a custom-built video player that displays concepts on a timeline. Besides searching through concert recordings, users can also browse through interviews deploying automatically generated speech recognition transcripts.

ACKNOWLEDGEMENTS
“Pride of Holland at Pinkpop” received financial support from the Dutch national research programme MultimediaN, technology foundation STW, and the project Images for the Future.

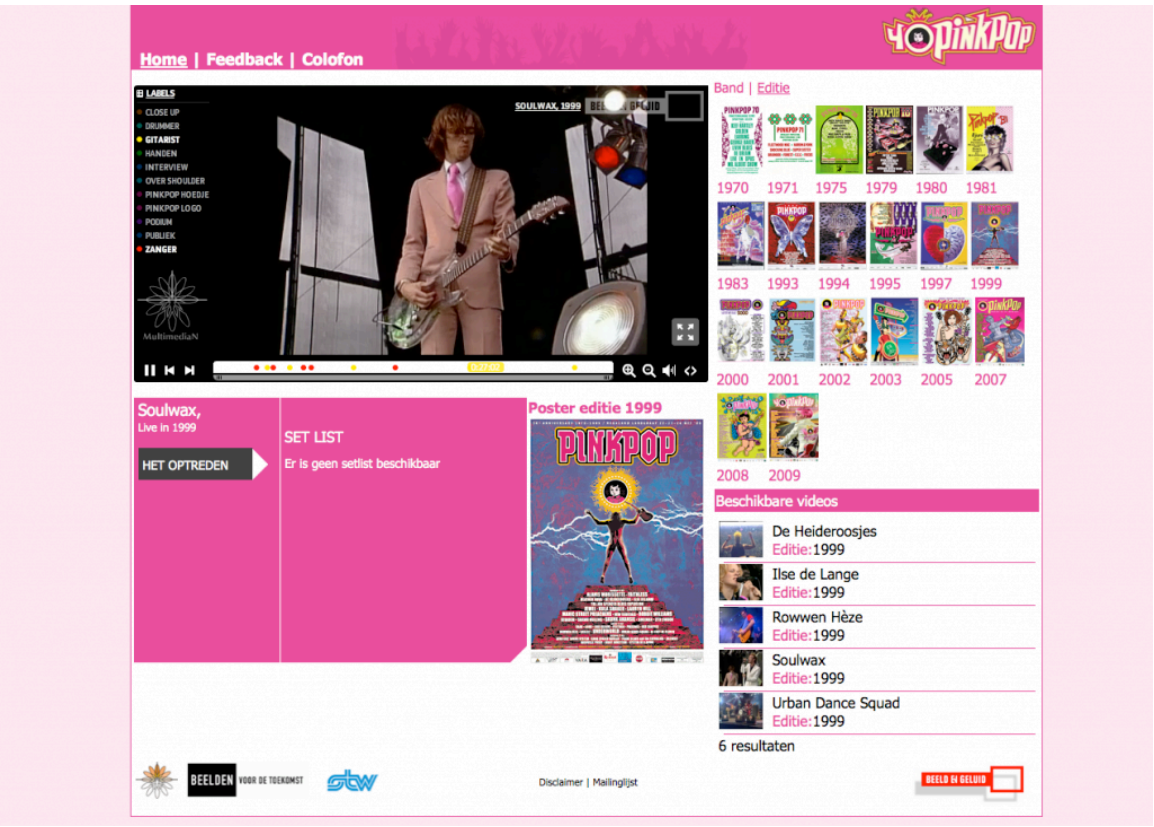


Fig. 1. Retrieval interface

Video Active and EUscreen
Providing Audiovisual Content to Europeana

Johan Oomen
Netherlands Institute for Sound and Vision
Sumatralaan 45, Hilversum
The Netherlands
+31-624151991
joomen@beeldengeluid.nl

Anna Christaki
National Technical University of Athens
28 Oktovriou (Patision) 42,
10682 Athens,
Greece
+30 210 7722017
achristaki@image.ntua.gr

Vassilis Tzouvaras
National Technical University of Athens 28 Oktovriou (Patision) 42,
10682 Athens,
Greece
+30 210 7722017
tzouvaras@image.ntua.gr

ABSTRACT
Many audiovisual archives are in the process of digitising their material and are exploring the new possibilities this brings to publish their content online. This paper provides insight into the background and development of the award winning Video Active Portal (thousands of video items are accessible through www.videoactive.eu) and its successor EUscreen (www.euscreen.eu); initiatives offering access to television heritage material from archives across Europe.

The Video Active project has used the latest advances in Semantic Web technologies in order to provide expressive representation of the metadata, mapping heterogeneous metadata schema in a common metadata schema based on Dublin Core, and advanced query services. As one of the main outcomes, the project successfully integrated their data to Europeana. The work of Video Active will be continued in the EUscreen Best Practice Network, launched in October 2009. In this three-year project, more fine-grained access to video objects is provided for, using the EBUCore Set of Metadata.

In this paper, the authors will firstly outline the work done in Video Active and will elaborate on the architecture of EUscreen that will expand the possibilities for connecting data across cultural heritage organizations in a meaningful way.

1. INTRODUCTION: ONLINE ACCESS TO AUDIOVISUAL HERITAGE
The greatest promise of the internet as a public knowledge repository is to create seamless access for anyone, anywhere, to all knowledge and cultural products ever produced by mankind. Mainly due to increased bandwidth availability and affordability of video cameras, web sites offering online video material have managed to mature and in a short period have become extremely popular. Web sites like YouTube, Vimeo, Facebook, Blip.tv and many others show how the idea of making and manipulating images (once mostly the preserve of professionals) has been embraced as a way of broadcasting who we are to anyone prepared to watch. The most popular site to date, YouTube, was launched in early 2005 and serves one billion videos per day [1]. The number of user generated video uploads per day is expected to go from 500,000 in 2007 to 4,800,000 in 2011 [2].

Looking at these numbers, it's evident that the potential for releasing material from audiovisual archives online is enormous. To date, however, from the many millions of hours in these archives a small percentage can be found online [3]. Although audiovisual content is now being digitised and some of it is available via the Internet, online access is mainly at the national institutional level, resulting in a wide range of conflicting and competing access routes. Many broadcasters in Europe have their own websites, with a growing collection of video items. Much of it is distributed across programme and subject-related 'pages'. There is little in the way of cataloguing for the bulk of material online, much less alignment with library standards for integrated access. Audiovisual archives need to overcome several obstacles before they can set up meaningful online services. These include: managing intellectual property rights, technological issues concerning digitisation and metadata standardisation and issues related to the way the sources are presented to users.

Recent advances in encoding standardisation and semantic web technologies and the widely adopted Dublin Core metadata element set have provided archives with the tools necessary to build meaningful services that provide unified access to archive videos online. A leading example in the audiovisual archive domain, the Video Active portal demonstrates that unified access to archives across different counties is possible. 14 of Europe's leading archives are publishing parts of their digitised assets on this platform, that is also linked to Europeana (www.europeana.eu); the European digital library. The project is supported by the eContentplus programme from the European Commission. The project started in 2006 and will run until September 2009.

2. VIDEO ACTIVE
Video Active has created a pool of television archive content (10.000 video's) and contextual data (articles, stills, program guides), representing national and cultural specificities of different European countries over a range of themes and historical events. Contributing archives include: BBC (UK), INA (FR), DR (D), DW (D), ORF (AT), NAVA (HU), Sound and Vision (NL) and many others.[1] The portal supports various textual search modes as

[1] For a complete list: http://videoactive.wordpress.com/the-consortium/

well as faceted, thematic and timeline-based browsing. (see Figure 1. And figure 2.)

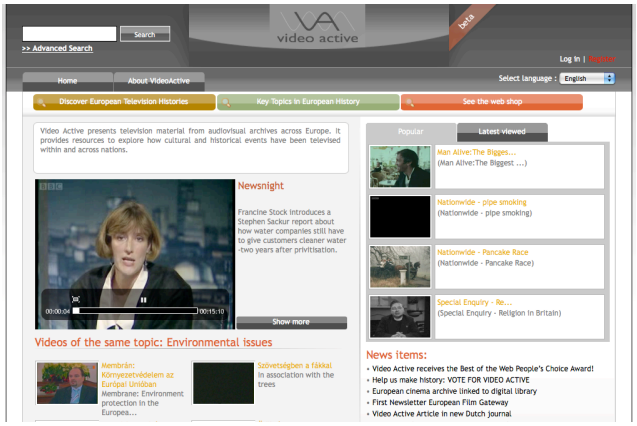


FIG 1. The Video Active homepage and results page

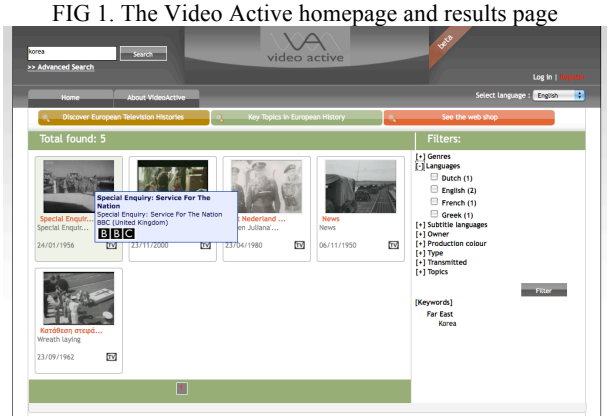


FIG 2. The Video Active results page

2.1 Defining requirements for Video Active

The demand for access to audiovisual content online has been growing, in a number of distinct sectors: education the general public and the heritage sector. For instance, digitization of archive content transforms cultural heritage into flexible 'learning objects' that can easily be integrated into today's teaching and learning strategies. These user groups have specific expectations and profiles, and the Video Active project had to understand and encompass these to ensure user satisfaction and revisits. Surveys, interviews and desk research have been executed in the initial stages of the project. The resulting insights in user requirements became fundamental to define the technical specifications and hence the technical architecture. Usability tests have been executed on the two consecutive releases of the portal. The excellence of the portal was acknowledged during the Museums and the Web conference 2009, where Video Active won the Best of the Web award.

2.2 The high level architecture:

The Video Active system comprises of various web modules. The whole workflow from annotating, uploading material, transcoding material, keyframe extraction, metadata storage and searching is managed by these components. Figure 2 shows the architecture of the Video Active portal. The architecture exploits semantic web technologies enabling automation, intelligent query services and semantic interoperability with other heterogeneous digital archives. In particular, a semantic layer has been added through the representation of its metadata in Resource Description Framework (RDF). The expressive power of RDF enables light reasoning services, merging/aligning metadata from heterogeneous sources and sophisticated query facility based on SPARQL RDF query language. Additionally Relational database and full-text search technologies have been used to store data where semantic information is not required and to improve the quering performance of the overall system. Finally, the Video Active metadata is public and ready to be harvested using the OAI Metadata Harvesting Protocol.

2.3 Storing and querying data the semantic way

The Video Active metadata schema is based on the Dublin Core set of metadata schema with additional elements (i.e. Genre, English Title) essential in capturing the specific properties of the resources [4]. The video metadata are generated automatically and are represented in a schema that is based on MPEG-7. In order to enable semantic services, the metadata is transformed in RDF triples and stored in a semantic metadata repository.

The annotation process is either manual or semi-automatic. In the semi-automatic process, the archives export their metadata (the ones that have mappings to the Dublin Core elements) using a common XML schema. The elements that cannot be mapped to the Video Active schema (or are missing from the legacy databases) are inserted manually using the Web Annotation Tool. This tool allows entering and managing the metadata associated with the media and also handles the preparation of the actual content. It contains the Transcoding Factory module that transcodes the original format of the source material to Flash and Windows Media streaming formats, creates low and medium bit rates for the streaming service and performs keyframe extraction for thumbnail creation. The Web Annotation Tool produces an XML file that contains metadata, based on Dublin Core, as well as content encoding and key frame extraction information. The XML is then transformed into RDF triples and stored in the Sesame semantic repository. Sesame is an open source Java framework for storing, querying and reasoning with RDF [5]. It allows storing RDF triples in several storage systems (e.g. Sesame local repository, MySQL database). The use of an ontology language, such as RDF that has formal semantics enables rich representation and reasoning services that facilitates sophisticated query, automation of processes and semantic interoperability. Search and retrieval in Video Active is performed using a combination of structured RDF queries in SeRQL (optimization of SPARQL query language for Sesame) and full text search queries using the high-performance, full- text search engine library Lucene.

All metadata stored in Sesame are exposed to external systems/archives with the help of an OAI-PMH compliant repository. Europeana, bringing together hundreds of collections of resources throughout Europe, has already indexed the data from the Video Active repository. (see Figure 3.)

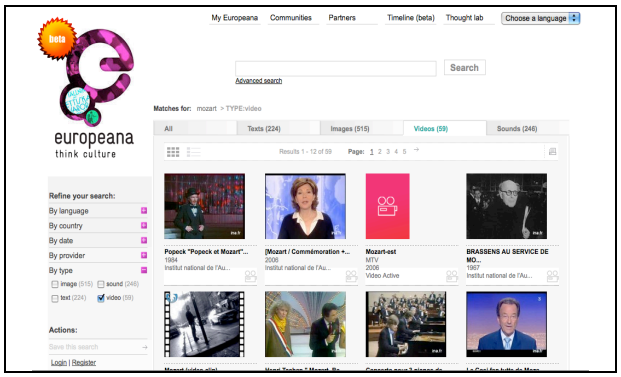


FIG 3. Video Active metadata indexed by Europeana

In order to exchange both the structure and the semantics of the metadata in a machine understandable way, distributed OWL/RDF query mechanisms will be employed in a future release.

2.4. Multilingual access: using SKOS

Eleven languages are supported in Video Active, this number will grow to 15 within EUscreen. The Video Active portal offers language support in four distinct ways. Firstly, Video Active has localized interfaces for each of the languages covered. Secondly, key metadata elements (i.e. DC Title, DC Description) are translated in English and thus provide the platform with a monolingual baseline. Thirdly, Video Active is using multilingual controlled vocabularies for the metadata elements Keywords, Genre and Location. The thesaurus from the International Press and Telecom Council is used as baseline for the keyword vocabulary. This 1500-term thesaurus has been translated by the Video Active project in 11 languages. For genres, the ESCORT 2007 EBU System of Classification of Radio and Television Programmes is used and for geographical names, the ISO 3166 English Country Names and Code Elements. Handling the translation of these terms and the export of these terms to machine readable XML is done in a specialized application called ThesauriX [6]. In order to achieve semantic interoperability the thesaurus taxonomy has been transformed into a semantic web language using the Simple Knowledge Organisation System (SKOS) standard. SKOS is a recommendation of the World Wide Web Consortium for the representation of thesaurus taxonomies. The SKOS standard is built on top of the RDF language and can be used to facilitate semantic retrieval of metadata and thesaurus alignment. Finally, a timeline view provides a visual overview over the milestones in the development of television in Europe using the SIMILE framework [7].

3. EUSCREEN: INTRODUCING LINKED DATA AND EBU CORE

The EUscreen Best Practice Network brings together 28 institutions (including 20 archives) from across Europe. To a large extent, EUscreen follows the same workflow as Video Active. Much of the software components will be reused. One of the differences is the fact that EUscreen adopts the EBU Core set of metadata that has been identified as being the minimum information needed to describe radio and television content [8]. An XML representation is provided in case this metadata would be implemented in archive exchange projects using the Open Archive Initiative's Protocol for Metadata Harvesting. Also, the integration with Europeana will change. Since Europeana and EUscreen will have their metadata represented in OWL/RDF and stored in semantic stores, a SPARQL end-point can be created to enable remote access. This way, Europeana will have the ability to directly perform SPARQL queries to the EUscreen repository (and vice versa).

Furthermore, EUscreen will deliver a number of additional web services. EUscreen will focus on implementing advanced reasoning and query services on top of the Semantic Repository (see section 2.3) to enable metadata post-processing such as consistency checking and extraction of implicit knowledge. Automatic enrichment will be investigated, enabling metadata enrichment with complementary data, information, and knowledge from external web resources (e.g. Getty, UNESCO thesauri, Wikipedia) available as so-called Linked Data, using URIs and RDF [8]. Also, using the metadata enrichment process, geographical coordinates will be assigned to all geographical places in order to visualize the results using the Google Maps API. Social networking functionalities will be developed, exploiting the recent advancements in Web 2.0 technologies in order to employ a framework where users can create their own galleries and participate in groups of interests, and where users can comment, annotate, tag, recommend or link up content items for personal or community use. Finally, the EUScreen metadata will be exported in other metadata schemas and in various language formats aiming at making it available in e-learning, leisure and research application scenarios.

4. ACKNOWLEDGMENTS

EUscreen and Video Active are funded within the eContentPlus programme of the EU.

REFERENCES

[1] Context is King: How Videos Are Found And Consumed Online. (2010) http://www.techcrunch.com/2010/01/30/context-is-king-how-videos-found/

[2] Manjoo, Farhad (2009) Do You Think Bandwidth Grows on Trees? Retrieved April 20, 2009, from: http://slate.com/id/2216162

[3] Wright, Richard (2007) Annual Report on Preservation Issues for European Audiovisual Collections. Retrieved April 20, 2009, from: http://www.prestospace.org/project/deliverables/D22-8.pdf

[4] Venetis, Tassos, Anna Christaki and Vassilis Tzouvaras (2007) Video Active: Domain and Upper Harmonizing Ontology. Retrieved April 20, 2009, from http://videoactive.wordpress.com/press/

[5] Broekstra, J., Kampman, A., Harmelen, F. (2002) Sesame: A Generic Architecture for Storing and Querying RDF and RDF Schema. In: 1st International Semantic Web Conference, Sardinia, Italy

[6] Gerhard Janisch. (2008) Analyse von Rich Internet Application Frameworks am Beispiel einer Thesaurusverwaltung, Joanneum Research, Graz

[7] Alonso, Omar, Gertz, Michael and Baeza-Yates, Ricardo (2007): Search results using timeline visualizations. In: Proceedings of the 30th Annual International ACM SIGIR

Conference on Research and Development in Information Retrieval 2007. p. 908

[8] EBU – TECH 3293-2008 (2008). Core Metadata Set for Archives (EBUCore) Specification v.1.0. Retrieved April 20, 2009, from: http://tech.ebu.ch/lang/en/MetadataSpecifications

[8] Tim Berners-Lee (2006) Linked Data. http://www.w3.org/DesignIssues/LinkedData.html

POSTERS

Mobile Video as an Instrument in a Social TV Service

Minna Suovirta
VTT Technical Research Centre of Finland
P.O. Box 1000
02044 VTT, Finland
+358 20 722 5916
minna.suovirta@vtt.fi

Markus Ylikerälä
VTT Technical Research Centre of Finland
P.O. Box 1000
02044 VTT, Finland
+358 20 722 6082
markus.ylikerala@vtt.fi

Chengyuan Peng
VTT Technical Research Centre of Finland
P.O. Box 1000
02044 VTT, Finland
+358 20 722 6839
chengyuan.peng@vtt.fi

ABSTRACT

Video derives as a larger part of social media its value from active participation of services with opportunities for social interaction and engagement. Into social video services can be integrate other communication services like voice, chat, context awareness, and ratings to support a shared TV experience within a community. We make a research and pilots to find out how mobile video can effectively be utilized in service provision and increase mobile TV experience.

In this paper we describe our vision and a platform of a mobile social video service. We show with three pilots the social video concept which can provide a great opportunity for content providers to keep directly in touch with users. The system also uses metadata based on semantic web for social discovery of video content generated by mobile users and for adapt content to what users want.

Categories and Subject Descriptors

H.1.1 **[Models and Principles]**: Systems and Information Theory – *Value of information*. H.3.3 **[Information Storage and Retrieval]**: Information Search and Retrieval. H3.4 **[Information Storage and Retrieval]**: Systems and Software – *Information networks*. H.3.5 **[Information Storage and Retrieval]**: Online information services. I.2.7 **[Artificial Intelligence]**: Natural Language Processing – *Speech recognition and synthesis*.

General Terms

Performance, Reliability, Experimentation

Keywords

Mobile Social Video, user generated content, metadata

1. INTRODUCTION

The television industry is facing one of its biggest changes in the form of personal computers, broadband Internet, and mobile. There is an on-going rapid expansion of video contents on the Internet and progress towards the "3-screen TV" model in reference to the growth of video on both PC and on mobile devices. There is plenty of room to expand the social interaction and participation opportunities for video content. This calls for innovation in developing novel social television services. There will also be plenty of possibilities in new mobile services, including those users who will become co-creators. The mobile social video will be integrated and interact in complex ways with other forms of video use.

Professional content providers and individual consumers will work in new interactive co-operation in the future of social video. The new systems like social video will help individuals and communities to publish fast and good quality video content with other individuals but also in co-operation with professionals. The roles of the players are changing so that consumers will be important content providers and professional companies will have new responsibilities and also new business opportunities.

Co-creative video production activities will be based on the collective power of many creators and call for new solutions on media metadata, real-time annotation and cross-linking of media sources, automatic media composition, innovative media formats, and new solutions for handling of intellectual property. This means call for novel tools and services that provide easy tools for capture and sharing, context-awareness in video creation, support for co-creativity among users, and direct integration into media production and distribution infrastructure.

Our social video concept is general and independent of industrial domain and it boosts services for third parties, whose business may have nothing to do with video itself. We are considering both content and technology. Content can be everything from mass media by professional creators to user generated media. Similarly technology aspects are not tied to a certain network or distribution mechanism. However, the focus is on video produced by individuals and communities in co-operation with professionals.

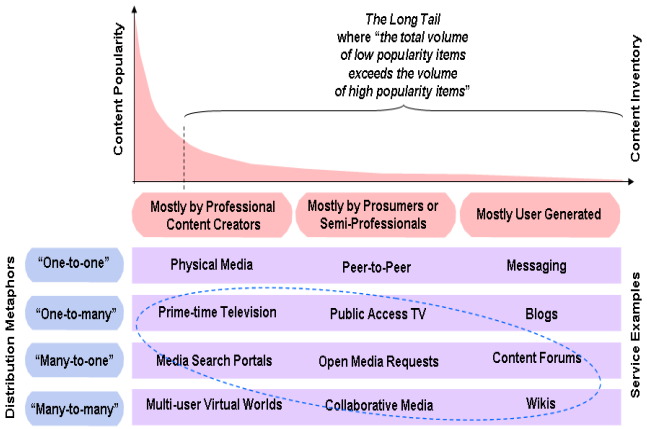


Figure 1. Scope of the project

Scope of the project [1] is shown in Figure 1. The base to social video service concept development is to find out (user and other) requirements for next generation collaborative social video production. We research and develop both technology and user related aspects.

In a social video service should be something that motivates users to use and contribute videos. Project partners'[1] knowledge and technologies are used to make an integrated social video concept in the multi channel environment (cableTV, IPTV, mobileTV, localTV) with hopes to offer good user experience.

In order to individual videos to be usable we need some ways to select, validate and package those. Our main point in this field is to identify false information from user generated content and select, filter and collect information from different sources to useful packages for different social video usages.

In the service concept development common standards for meta-data and interfaces play important role in successful results. The current major problem of the semantic search is the lack of available semantics for the resources. Users should have different easy ways to inject keywords to be used as metadata.

2. ARCHITECTURE AND PLATFORM

In this section we present our system and platform architecture.

The SocialVideo Distilled is a location based service (LBS) that passes videos into possession of location for collectively improving available videos with user generated metadata. Figure 2 shows the architecture of the project that basically consists of video and metadata management from heterogeneous sources, and tracking of the user based on the user's mobile device. The user interface on the mobile device can be used for user based metadata generation and broader delivery of the improved distilled videos.

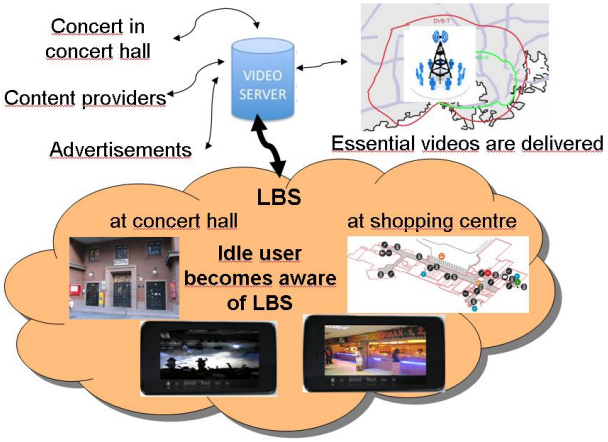


Figure 2. Social Video Distilled as LBS

The uSpace [2] provides a platform for the actual service i.e. SocialVideo Distilled. Figure 3 describes the mobile platform building blocks based on uSpace. The uSpace evolved within LUCRE [3] service development kits is a collaborative and ubiquitous service engine. It extends current concept of widget containers, e.g. iGoogle and Netvibes. End-users can create and manage dynamic web services combining a set of widgets from a library to a shared space. It also aims for spatially and temporally aware services which treat location and time as metadata.

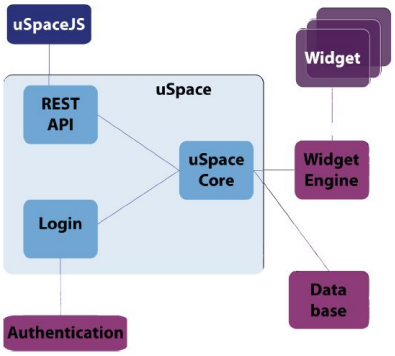


Figure 3. Architecture of the uSpace and the relating components [2].

Thus, a logical video server contains arbitrary multivendor videos such as from concerts as well as from advertisers or combined with advertisements. A user on the location (e.g. the concert hall) becomes aware of the service based on the tracking of user's mobile device by e.g. GPS, WLAN, Bluetooth, etc. The user can further access the service with the mobile device. Next, the service provides the previously recorded videos at the concert hall for the user to comment, grade etc and the user's contribution is finally included to the service.

3. METADATA USAGE

An essential aspect of these approaches is the inclusion of metadata. Metadata is essential for helping recommendations, search and discovery. The service can keep control of the tags and metadata and link users to content via a specific event. Semantic search of content could aggregate video from several sources, personalize them to the user interests and location, and to augment the probabilities for content to be found faster and with more relevance to users.

We designed an end-to-end metadata architecture for semantic and collaborative search of content shown in Figure 4. The central part is the deployment of Drupal content management framework [4]. The content metadata (knowledge) store is a structured collection of records of metadata. We use the Resource Description Framework-in-attributes (RDFa) as metadata format [5]. This architecture is responsible for serving xhtml web pages with embedded RDFa metadata queried by uSpace platform. This usage architecture can build a distributed and collaborative editorial model in broadcaster site.

The modules in Figure 4 (in next page) including the third party modules are functional plug-ins that can interact with the dynamic content databases created by end users. The process of configuring RDFa support is simple and straightforward.

[1] Arcada University of Applied Science, Digita Oy, Elisa Oyj, Forum Virium Helsinki, Helsinki Institute for Information Technology (HIIT), Lingsoft Oy, Nokia Oyj, Sofia Digital Oy, Sanoma Entertainment Oy, Sanoma Television Oy, Teosto ry, University of Tampere, Tampere University of Technology, TVkaista Oy, Vincit Oy, VTT Technical Research Center of Finland

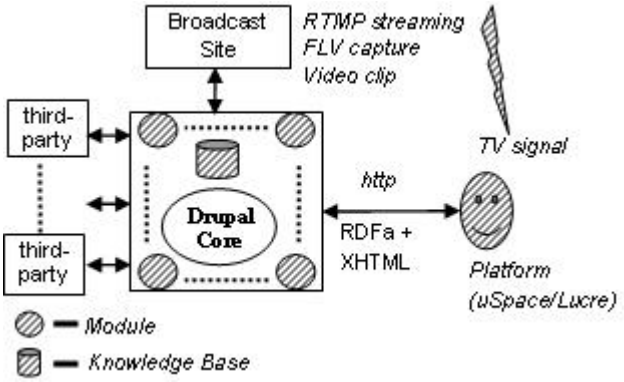


Figure 4. Metadata Flow Diagram.

In the project we use the concept of metadata bundle, that is, in addition to the base video metadata element set (cf. table 2), other optional metadata sets can be combined for different scenarios (cf. table 3, contextual metadata elements).

Table 2. Base Metadata Element

Attribute	Vocabulary	Description
title	DCMES	title of the video file
creator	DCMES	creator of the video file
date	DCMES	date that the video was uploaded
identifier	DCMES	URL of the video file
subject	DCMES	a list of keywords describing the content of the video
format	DCMES	format of the video file
rights	DCMES	a URL to a text-based description of the video's licensing terms
description	DCMES	provides a text-based description of the video
relation	DCMES	relationship of this resource to another resource
source	DCMES	objects, either electronic or printed, from which this resource was derived
type	DCMES	the nature or genre of the content of the resource
extent	DCMI	the size or duration of the resource

The Simple DCMES (Dublin Core Metadata Element Set) consists of 15 metadata elements, from which we found main part suitable for us. The additional terms are identified, generally in working groups of the Dublin Core Metadata Initiative (DCMI).

Table 3. Contextual Metadata Elements

Attribute	Vocabulary	Description
position	GeoTag	latitude;longitude position coordinates for location information about the current location of the device
placename	GeoTag	the name of the place
region	GeoTag	the name of the region, country
weather, temperature, humidity	SensorML	Current environmental conditions including weather, temperature, and air conditions
atDateTime/created	Timeline Ontology /DCMI	the current system time of the content consumption

4. PILOTS

We have three pilots in different areas. There are many interesting questions to which we hope to get useful answers via our pilots. We among others think issues relating metadata, user-generated content, editing and users of the video content.

Users of mobile phones are beforehand recruited and advised to use a test mobile phone with appropriate applications. User tests and interviews are performed both during the development process and in the pilots in order to contribute social service development.

4.1 Live concert

We have couple of live concert user tests to get needed mass of people and variation of concert situations to get reliable user feedback results. We use both an automatic and a manual remix, so that we can compare implications and the social activity the different editions create. The concert pilots are arranged on April 2010.

In a band gig several fans record video clips by mobile phone. After the gig fans upload their clips into a social video service. A remix video, made from automatically selected video clips, is played on music club's widescreen right after the gig. The video clips are also put available to fans for manual remixing. Both editions are put available for the fans to view and discuss in the band's web page.

In a music hall are series of band competition concerts. Concerts will be recorded both by mobile phone cameras and professional equipments. Mobile phone shooters change in the concerts so that we get feedback from various test users. Mobile phone users upload their video clips into a server and an automatic video clip remix is also done in that case so that broadcast editors can use it in remixing professional multi-camera material and user-generated material into a live video entity. In that way we get results how professional broadcast content can be enhanced with user generated content during a live event. After the edition the concert materials are delivered in multiple channels to test multi channel environment.

To the social video service can be included speech recognition as a way to users to give and system to find out metadata. Speech recognition is much easier from video clips taken in a less noisy event, so we test speech recognition as a source of metadata partly separately to above mentioned pilot cases.

In both cases it can be seen that all interested parties (the band/, band's fans and the music club/hall, etc.) have their own benefits and interest to promote the social video production.

4.2 Added value broadcast

Traditionally friends, families and other communities get together and gather around television to watch e.g. live sports event or a TV series. They discuss and communicate with each other at least during the event or an episode.

The social aspect of gathering physically together can be transferred to web environment by allowing commenting and sharing video messages and synchronized feed with each others during a live event. Other synchronized video material is streamed to the community (an open/enclosed auditorium in Facebook), no matter where the participants physically are as long as they have the internet connection. The communal and social video viewing

is transferred to a web site by means of a computer or mobile phone.

These communities extend far beyond the home to span entire neighborhoods, cities, countries, and hemispheres. And like the traditional living room, these communities are increasingly organized around video, connecting families, friends, and some strangers alike in a shared video space defined by interactions, common interest, or location.

Two sessions of the pilot are arranged on April. Test users gather together to view same sport broadcast and use system to communicate each other. The pilot tells us, how well social commenting possibilities and user generated contents bring added value to the professional broadcast during a live event.

4.3 Case of Stories

Group of students are given a task to record five minutes of their life every day for a week concentrating on a given theme. The students can include in beginning of the video clips some optional predefined keywords e.g. 'place school' (in Finnish) to be detected by speech recognition. Pilot tests are arranged in beginning of May. The resulting videos are tagged, archived, and edited into a story which can be used among other content in delivery of multiple channels.

The goal of the pilot is to show that user generated social content can make a big story. The case is also part of StadiTV [6] project.

5. CONCLUSIONS

There are many drivers to use a video as a social media instrument. Users are used to take videos by mobile phones and video cameras, and send clips to relatives and friends. Why don't send conveniently and fast a descriptive video than write same message e.g. in a letter?

There is huge potential to user generated content creation in itself and as enhancement of a professional media. The professional media actors has pressures to utilize user generated content and expectations social video to be a new way to engage and motivate users to be their customers.

A video presentation may tell much, but a file name automatically generated by a digital camera does not. To search, find, group, etc. video materials you need to have metadata with video files. This calls for good video services to save materials.

A video service could find automatically part of the basic metadata e.g. time, but all significant metadata can not be got automatically and a machine may misinterpret. So, the users should get motivated to add metadata. One possibility is to use automatic speech recognition in giving and finding out metadata.

We test speech recognition in our pilots in giving metadata and taking part in a social session within a video. We expect good results, if background noise is not too loud. After the pilots we can say more about acceptance of the speech recognition feature among the test users.

Project has got good results in automatic detection of e.g. time and location metadata from video data, which is created within the project or linked e.g. from YuoTube. We are also able to substantially identify false information and cluster videos to useful packages for different social video usages.

In the pilot cases are various motivating issues for users to use and contribute social videos in the service. After the pilots we can say more about motivation issues.

The user generated content as well as professional one should be edited. Everybody does not want, has time or can do that. So, it would be nice to put own clip to a service and in addition to see others' clips e.g. from the same concert and get even a remix of whole event. After the pilots we can say more about successfulness of that scenario.

We deliver video content within copyrights in multiple channels. After the pilots we can say more about successfulness of multi channel delivery and utilization of same materials. Copyrights affects to it, what is possible and allowed to do with individual video materials. Material authors (composers, bands, etc.) put it own challenges to utilization of materials. Copyrights to a one-time-use in e.g. a concert hall's wide screen can be quite easily arranged, but how about in case of e.g. on-demand multi-channel delivery? User generated content and their materials delivery via many different channels challenges copyright organizations to license repertoire development and further international collaboration between actors in different countries.

Current pilots are our ongoing work. We will present detailed results after. Our main contribution is the design of a video-oriented mobile social TV service system in which various social features can be technically integrated with the actual mobile TV viewing experience.

Current mobile TV services could certainly benefit from this augmentation - from the traditional uni-directional viewing screen to a more connected and shared space with family, friends, and communities. Semantic technologies used also make it possible to give richer descriptions to mobile video, facilitating the process of locating, combining diverse media from various sources and personalizing content recommendation.

6. Acknowledgement

The authors would like to thank project participants for their inspiriting visions and fruitful work to develop a concept and platform for a mobile social video service. Special thanks to Arcada, Forum Virium, HIIT and Sanoma for opening up good pilot case topics.

7. REFERENCES

[1] Social Video, http://www.flexibleservices.fi/en/node/32.

[2] uSpace, http://www.flexibleservices.fi/files/file/pdf/LUCRE_D25_ReleasesOfLUCREToolsAndSystem.pdf

[3] LUCRE (Local User Created Services), http://www.flexibleservices.fi/en/node/28

[4] Drupal, http://drupal.org

[5] RDFa, http://www.w3.org/TR/rdfa-syntax/

[6] StadiTV, http://www.forumvirium.fi/stadi.tv-2.html

Understanding users of web-TV

Amela Karahasanović[1,2]
[1]SINTEF ICT, P.O.Box 124 – Blindern, 0314 Oslo, Norway
[2]Department of Informatics, University of Oslo, Norway
+47 48 10 88 95
amela@sintef.no

Jan Heim
SINTEF ICT, P.O.Box 124 – Blindern, 0314 Oslo, Norway
+47 22 06 75 69
Jan.Heim@sintef.no

ABSTRACT
Understanding users' behaviour and expectations is a prerequisite for providing services tailored to their needs. We conducted a survey with 2,400 users to determine when, why and how they used a web-TV service and whether they were satisfied with it. In this paper, we present the preliminary results of this study. Our results show that the major motivations for using this service were related to content and flexibility. Furthermore, when describing their satisfaction with the service, the users took into account the effects of the equipment and network connection on their experience.

Keywords
Web-TV, User studies, Motivation, User Experience

1. INTRODUCTION AND BACKGROUND
The usage of web-TV and other multimedia services delivered over telecommunication networks is becoming increasingly widespread. With triple-play services (IPTV, IP-telephony and Internet), households have all their services delivered through a single broadband access link. Household members expect all of their services to work perfectly all the time, in spite of technical limitations. The R2D2 Networks project aims to make this possible by developing user-dependent, self-adaptive networks [1]. Understanding users' behaviour and perception of the services they consume is a first step in the development of such networks.

Everyday practices related to traditional TV, as well as experiences with video-on-demand systems over IP [2], watching football on web-TV [3], usage of mobile TV [4, 5] and experiences with IPTV [6,7] have been reported in the literature. Our work adds to these findings by exploring the usage of web-TV services.

2. METHOD
To understand users for the purpose of defining their requirements and evaluating their satisfaction when using a product/service, we need to answer the following questions:

- Who are the users? What are their demographic characteristics, preferences, expectations, skills and so on?
- What tasks are they going to perform? What type of media/equipment is going to be involved? In which context are they going to perform these tasks?
- What is their motivation and what are their expectations related to a product/service?

As a first step in our investigation of users of triple-play services in Norway and Europe, we conducted an explorative study with the following objectives:

- To obtain an overview of the behaviour patterns of users of web-TV in Norway.
- To test our questionnaire and identify the categories and expressions employed by users themselves when describing their experience with web-TV services.

In collaboration with a Norwegian web-TV provider, which will be called TV-X throughout the document, we developed an open-ended questionnaire. This provider broadcasts a wide spectrum of TV programmes, including news, sports, documentaries, TV shows and movies. Some of the content is broadcast exclusively on web-TV. The questionnaire (in Norwegian) was based on our previous work [8] and input from TV-X, and included questions about users, their utilization of TV-X's service and their experience with it. The questionnaire was created using SurveyMonkey and posted on the TV-X web site. After 25 hours, 2,408 users had answered and we stopped data collection. The full list of the questions, translated into English by the authors, is given in Appendix A. The results were analysed as follows. Firstly, we imported the results into an Excel file. We then analysed the 150 longest answers for each question (as measured by the number of characters) and identified the categories most often mentioned by the respondents. This was done by two researchers in parallel. Based on the categories identified, all 2,408 answers were then coded by a researcher. Basic statistics were produced in SPSS. The results were then presented to the web-TV provider. Information on the content provided, subscription fees and user satisfaction is confidential and is not presented here.

3. RESULTS
3.1 Users and Context
Most of the users of this service were young people between 18 and 26 years of age (Figure 1). The largest group of users had graduated from high school, while college graduates represented the second-largest group. Both genders were equally represented among the users. This is in line with the findings on gender and age reported in [9].

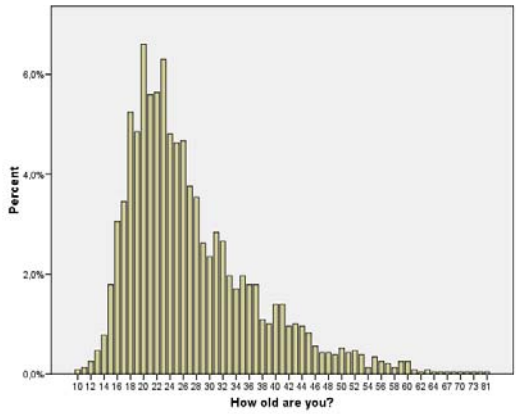


Figure 1. Distribution of age: 2,408 respondents.

More than 90% of the users watched web-TV at home. Watching web-TV at a job or at school was rarely reported. The majority of users watched web-TV at their computers, and very few reported the usage of extra equipment. TV sets were used by 35 (of 2,408) users, external screens were used by seven users and projectors were used by three users.

3.2 Tasks and Motivation

We asked the users about their intended activities (tasks) when they logged in to use TV-X (Figure 2). We also asked them why they subscribed to this service (Figure 3). To some degree, these answers overlap for the following two reasons:

- Some of the users answered the first question on TV-X ('Why are you using the service **now**?') by saying why they needed the service in general. For example, some of them mentioned that they do not have a TV.
- At the time of answering our survey, some users had logged in to watch their favourite program or because someone else in the family was watching TV, and this was also the main reason they had subscribed in the first place.

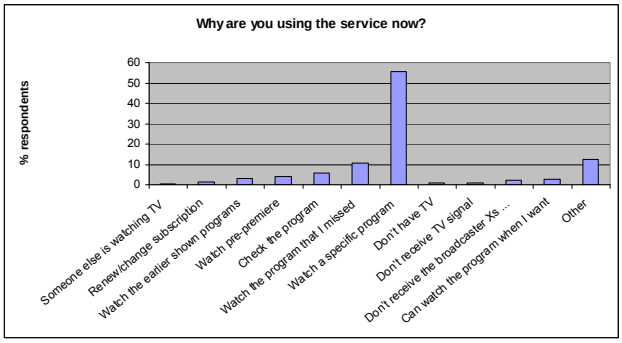


Figure 2. Reasons for using the service.

The most important reason for using the service (48%) was watching a specific, very popular soap opera. The next most important reason (10%) was watching a program that a user had missed (not necessarily the same programme that was given as the most important reason for using the service). All other activities, like checking a programme or watching a pre-premiere or programs that had been shown earlier, were reported by less than 5% of users. These results demonstrate the importance of the content in the context of web-TV. It is crucial to consider content and its different features, both in the requirement analysis and in the evaluation of user experience.

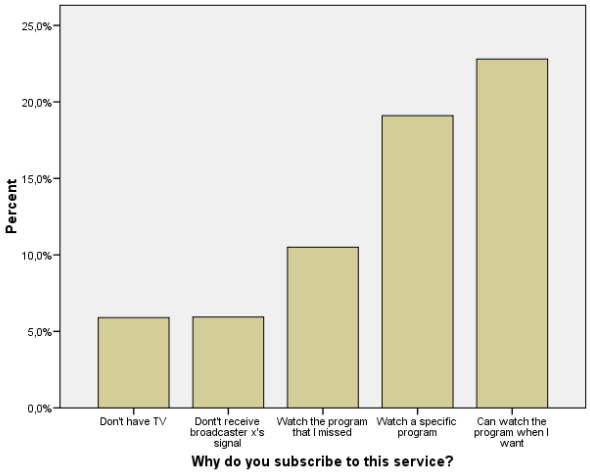


Figure 3. Main reasons for subscription.

The majority of users reported flexibility (watching the program when they want) as the main reason for subscribing to this service. The second and the third most important reasons were related to content (watching a specific program and watching a missed program). These were followed by practical/technical reasons (don't have TV and don't have signal).

3.3 User Satisfaction and Preferences

We asked several questions about user satisfaction and preferences. These results, however, are not presented here because, in addition to being confidential, they were also too specific. A result that might be of wider interest is the difference between user satisfaction with direct programming and earlier prepared videos. Figure 4 presents these results. The majority of users were satisfied with the service and, unsurprisingly, they were more satisfied with the quality of the video than with the quality of the direct programs on webTV. We found no correlation between the browser used and user satisfaction. Furthermore, some users commented that they were very satisfied even though they were using poor quality equipment.

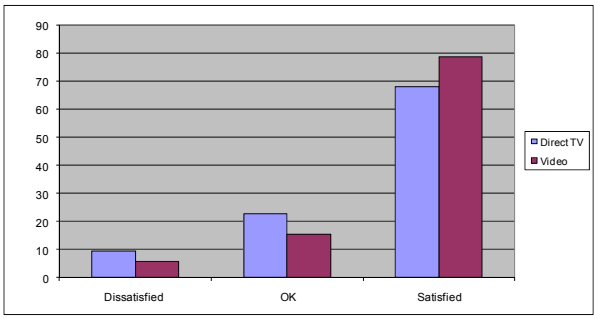


Figure 4. Satisfaction with the sound and picture when watching direct TV/ video.

Other users reported that they were not satisfied, but that they assumed that the problem had to do with the limitations of their network connection. An important lesson learned from this is

that the evaluation of complex products such as web-TV should take into account the responsibility of different providers as perceived by users.

4. CONCLUSIONS AND FUTURE WORK

In this paper, we presented preliminary results of a survey we conducted with over 2,400 users of a web-TV service. The results show that the service was most popular among young people who liked to watch their favourite program at home. Flexibility and content (watching a specific program when initially shown or later on) were the most important motivating factors for using the service. This is in line with similar studies of IPTV [7]. Considering content is thus very important, both in the requirement analysis and in the evaluation of user experience. An important lesson learned is that users are aware of product complexity and the effects of their equipment or network connection on their experience of the web-service. We plan to investigate this issue further.

The sample we used was a convenience sample. The users who were most interested in improving the service answered the survey and the validity of our results is therefore limited. We plan to conduct a similar survey with users of another web-TV service, along with a survey of a representative sample of Internet users in Norway. This will be complemented with studies analysing network traffic data and log data on the service providers' servers, as well as studies conducted in other European countries.

5. APPENDIX A

About you

1. How old are you?

2. Gender (M/F)

3. Where do you live?

4. Do you have a degree? Are you a pupil or student? If so, please specify the degree you are currently pursuing.

5. What is your household's gross income per year (before taxes and tax relief)?

About TV-X

6. Why are you on TV-X now?

7. Why have you subscribed to TV-X?

8. What kind of subscription do you have?

9. Who is paying for the subscription?

10. Where are you now? Home, work, school?

11. What kind of equipment are you using now? PC, laptop?

You opinion about TV-X

12. How satisfied are you with TV-X?

13. What do you like most?

14. What should be improved? Is there anything that you would like to see provided?

15. How satisfied are you with the sound and picture when watching direct programme on web-TV?

16. How satisfied are you with the sound and picture when watching video?

Finally

17. Do you have any comments on this survey?

6. ACKNOWLEDGMENTS

This work was supported by the R2D2 Networks research project funded by the Norwegian Research Council VERDIKT Programme (project number 193018/S19). We thank our industrial partners in the project as well as the participants in this study.

7. REFERENCES

[1] Areizaga, E. Perez, L. Verikoukis, C. Zorba, N. Jacob, E. Odling, P. 2009. A Road to media-aware user-Dependent self-aDaptive NETWORKS. In the proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, BMSB '09. IEEE. pp 1-6.

[2] Yu. H., Zheng, D., Zhao, B.Y., Zheng, W. 2006. Understanding User Behaviour in Large-Scale Video-on-Demand Systems. In the proceedings of EuroSys 2006. ACM Press. pp. 333-344.

[3] Raasakka, K.A. 2009. Can web-TV become a disruptive technology to television in broadcasting of Norwegian elite series football?, MSc theses, BI Norwegian School of Management, 2009.

[4] Miyauchi, K., Sugahara, T., Oda, H. 2008. Relax or Study: A Qualitative User Study on the Usage of Mobile TV and Video. EuroITV 2008, LNCS 5066, Springer-Verlag, Berlin, pp. 128-132.

[5] Vangenck, M., Jacobs, A., Lievens, B., Vanhengel, E., Pierson, J. 2008. Does Mobile Television Challenge the Dimension if Viewing Television? An Explorative Research on Time, Place and Social Context of the Use of Mobile Television Content. EuroITV 2008, LNCS 5066, Springer-Verlag, Berlin, pp. 122-127.

[6] Obrist, M., Kepplinger, S., Beck, E., Tscheligi, M., Muzak, P. 2008. Usability and User Experience: preliminary Results from Evaluating an IPTV Community Platform. EuroITV 2008, LNCS 5066, Springer-Verlag, Berlin, pp. 303-307.

[7] Dong, H.S. 2006. Potential user factors driving adoption of IPTV. What are customers expecting from IPTV? Technological Forecasting and Social Change. Elsevier, pp. 1446-1464.

[8] Brandtzæg, P.B and Heim, J. 2009. Explaining Participation in Online Communities. (Eds) Brian Whitworth and Aldo de Moor. *Handbook of Research on Socio-Technical Design and Social Networking Systems*. Information Science Reference (formerly Idea Group Reference). New York, USA pp. 167-182.

[9] World Internet Project – International Report 2010, Press Release, USC Anneberg School, February 2010, available at http://www.worldinternetproject.net/, accessed 6.4.2010.

Television Usability Tests with Conventional Remote Control and Wii Remote Using Prototypes

Christian P. Brackmann
Universidade Federal de Santa Maria (UFSM)
Colégio Politécnico da UFSM
Santa Maria / RS - CEP 97105-900
+55-55-3220-8070
christian@brackmann.com.br

Fabio R. Palamedi
Universidade Metodista de São Paulo (UMESP)
Curso de Comunicação Social - POSCOM
São Paulo, SP, Brasil
+55-11-9433-0061
fabio.palamedi@gmail.com

ABSTRACT
This research presents the state of art of the Brazilian Digital TV System (BDTVS) and statistics of the use of the television in citizen's house. The BDTVS were developed not just as a necessary evolution of the analog television, but also a tool for digital and social inclusion of the population. After developing several high-fidelity prototypes for both remote controls: conventional and Wii Remote (point-and-click), usability tests were applied to six different profiles compatible with the Brazilian population. This research also presents partial results and difficulties of the tests.

Categories and Subject Descriptors
D.2.M [Miscellaneous]: Rapid prototyping.

General Terms
Measurement, Design, Human Factors.

Keywords
Usability Tests, Digital TV, Prototyping

1. INTRODUCTION
Nowadays, the use of computing is not only restricted to computers. The microchip has a guaranteed place in the vast majority of electronics. It is present in phones, cameras, MP3 players (MPEG-1 Audio Layer 3), coffeemakers, alarm systems, and many others which we are not aware of, because they are used on natural basis.

Blending computers and telecommunications have made the television a high source of diversity. A century of fast television technological evolution has passed since the first invention but the information Era is just beginning; since the topic is digital television, certainly the most appropriate word is convergence.

Terrestrial TV diffusion occurs through electromagnetic waves (air) and is free. This characteristic shows the impact that this process of digitization will have on Brazil, because it is the one that has more televisions per household in the world. With the arrival of digital television across the country, broadcasters have had a big challenge to adapt the many ways of producing and distributing content (linear and nonlinear) of the latest technology for an audience of heterogeneous culture and social contrasts [2]. These viewers may be of various types, ages, regions, and social classes - in other words, the television content should be compatible with any kind of public.

Currently the interactive television is still not widely used due to technical, industrial, and governmental issues. The component which is responsible for its interactivity (Ginga [3]) has not been officially released. Its launch is planned for the year 2010, promising to bring many innovations, which will be available for the general public that has a digital signal (initially) and by the end of the year all brazilians will have access to this technology..

The government considers the Digital TV a great opportunity to become a great tool for social and digital inclusion for the poor, providing health services, social security and other functions in their homes. Table 1 demonstrates a small number of computers and Internet in households, however there is a high number of televisions. The vast majority of the Brazilian population does not use computers in their routine, and many do not even know how to handle a cell phone. What is the degree of difficulty of the interaction of the public "technologically illiterate" with interactive television? This question was the starting point of this research.

After studying the many difficulties that the population could have during their TV interaction, the Digital TV Interaction Research Group of the Pelotas Catholic University decided to test these users and discover through Usability Tests if these difficulties really apply.

Table 1: Television in Brazilian Households [4]

Regions	Percentage of households with some durable goods and services access to communication in the total of permanent private households (%)											
	Lighting power supply	Telephone Service		Stove	Water filtration (unit)	Refrigerator	Freezer	Washing machine	Radio	Television	Personal Computer	
		Total	Mobile								Total	Internet Access
Brazil	**98,6**	**82,1**	**37,6**	**98,2**	**51,6**	**92,1**	**16,0**	**41,5**	**88,9**	**95,1**	**31,2**	**23,8**
North	94,9	72,4	49,0	96,7	31,4	83,9	15,3	26,7	76,4	90,0	17,4	10,6
Northeast	97,0	66,8	43,9	96,1	49,1	81,5	7,0	15,5	82,4	91,7	15,7	11,6
South	99,8	88,9	29,3	99,4	65,0	97,3	15,6	54,3	93,0	97,6	40,0	31,5
Southeast	99,4	89,8	40,8	99,0	22,6	97,2	31,5	59,0	94,9	96,4	38,5	28,6
Midwest	99,2	87,9	47,7	98,8	60,4	95,1	18,6	34,4	86,4	94,6	30,9	26,5

2. Defining Profiles

In Brazil, Barros [1], after examining four researches about user modeling, made in the United Kingdom, the United States and Portugal, proposed in his paper **which** he thought a rating closer to reality for Brazilian population. However not all possible profiles were listed. After adding the missing profiles, the following classification of the User Templates were generated:

- **Juvenile:** children or teenager. Do not have any difficulties on interacting with the TV;
- **House Father/Mother**: adult, medium-low familiarity with technology. Interested on pre-defined TV shows;
- **Busy Mother**: adult, female, indifference to technology and attention dispersed among many activities;
- **On-Line Adult**: adult, good familiarity with technology. Watches television whenever possible;
- **Middle Age with Help**: adult, above 50 years. Needs glasses for short distances. Little familiarity with technology and asks for help;
- **Elderly**: adult above 65 years. Watches much television and does not like technology. Tends not to interact with television.

After defining the profiles that would be interviewed, it was necessary to develop prototypes compatible with both platforms: conventional remote control (tested on computers) and the Wii Remote (tested on the video-game consoles). The next section explains the prototyping technique and the purpose of each one.

3. Developing High Fidelity Prototypes

There were developed four prototypes for both control types and an extra one exclusively for the Wii Remote Control.

3.1 Common TV Series

Figure 1. User must click on the red trigger (Wii) or use the red button on the remote control (Conventional Remote).

3.2 Soccer Game

Figure 2. User should see the trigger contracted on the top-right corner and open the application

3.3 Big Brother Voting Poll

Figure 3. Big Brother voting application

3.4 Daily News

Figure 4. User should read at least one news on the interactive panel.

3.5 Change Volume and Channel (Wii Remote Only)

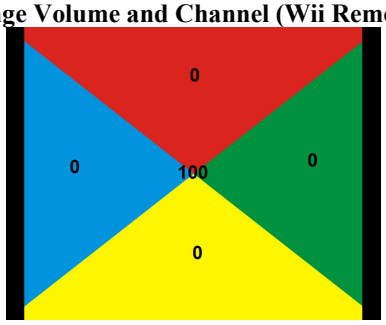


Figure 5. Using the Wii Remote to change the volume and channel without pressing buttons

The Wii Remote Movements for this test are available on Figure 8.

All the prototypes were developed using Adobe Flash 7, because the technology is capable of working in both machines: computer and Wii Video-Game Console through Opera Browser.

4. Usability Test Environments

There were used two environments, one for the Conventional Remote Control (Figure 6). Item A corresponds to recording camera, B is the sofa where the tester was sitten, C is the computer where the application was being executed and D is the infrared sensor. Figure 7 demonstrates the position of the equipment for the Wii Test. Note that the computer (Item C) is changed to a TV and the Wii Console was required for the tests (Item D). Letter E is the Wii's SensorBar (infrared lights) [5] [6].

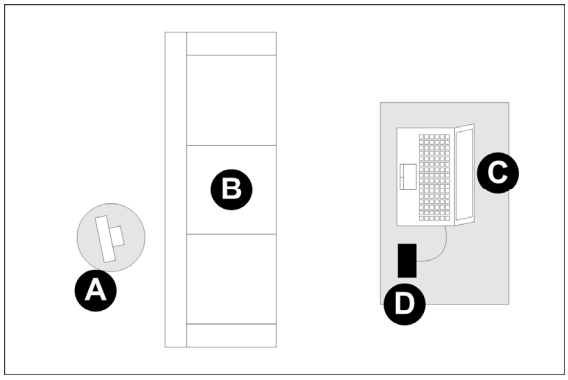


Figure 6. Test Environment for Conventional Remote Control

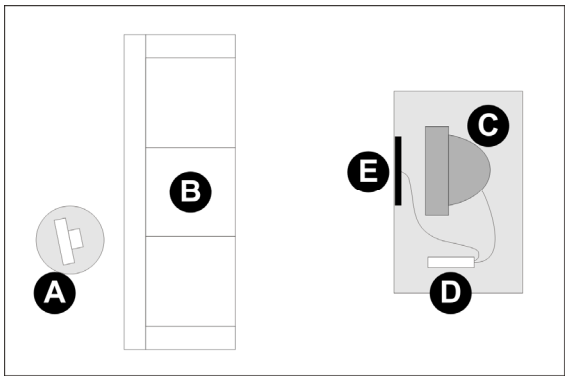


Figure 7. Test Environment for Wii Remote Control

After defining the prototypes, interviews points, and the necessary equipment, the testers started the research and made a notes of all difficulties from the participants.

The results can be found on the next section.

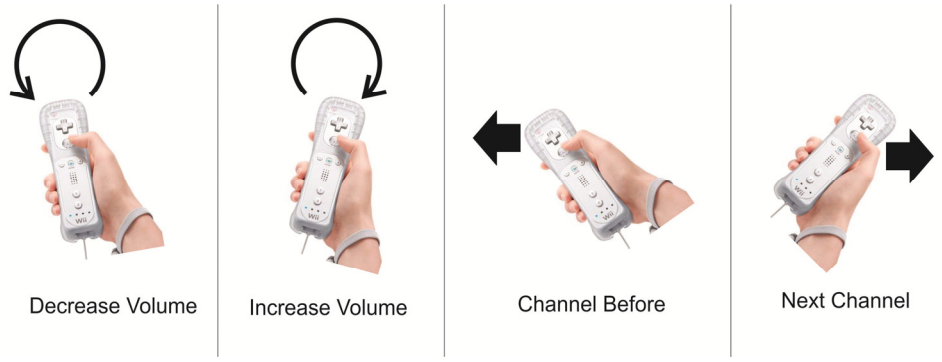


Figure 8. Wii Remote Movements

5. Results / Conclusions

- 67% of the participants didn't interact with the trigger;
- 11% didn't understand how to use the Wii Remote;
- The trigger should stay a long time on the screen. Elderly people take a long time to grab the remote control and interact;
- On-Screen Buttons should be in big proportions. Sometimes it's almost impossible to click on them with the Wii Remote;
- 17% of the participants had difficulties closing the application. The solution for the problem was turning the TV off and on again;
- 22% of the testers said that they were afraid of breaking their TV and lose their main entertainment equipment;
- If the user could read the news before the broadcasting they would still watch the News (linear stream);
- 56% of the interviewed still prefer the conventional remote control, because it is difficult to point to the TV and are not open to changing habits;
- The testers enjoyed both control paradigms, the best way would be that both versions could exist on interactive TV applications.

6. References

[1] BARROS, G. A consistência da interface com o usuário para a TV interativa. Universidade de São Paulo. São Paulo: POLI-USP. 2006.

[2] BECKER, V., & MONTEZ, C. TV digital interativa: conceitos, desafios e perspectivas para o Brasil. Florianópolis: UFSC, 2a ed. 2005.

[3] GINGA, Middleware Ginga. www.ginga.org.br. Access: 14/07/2009. 2009.

[4] PNAD2007, PESQUISA NACIONAL DE DOMICÍLIOS 2007 Relatório do PNAD 2007.

[5] NINTENDO. Nintendo of America. http://www.nintendo.com. Access em: 16/07/2009.

[6] PREECE, J; ROGERS, Y; SHARP, H. Design de Interação: Além da interação homem-computador. Porto Alegre: Bookman, 2005.

The need of user centred processes to do usable and accessible technology-design for DTV

Hans Persson
IHT / KTH
Bollnäs, Sweden
+46 278 258 67
hans.persson@iht.se

Francisco Fraile
Interactive TV Arena/ Universidad Politécnica de Valencia
Gävle, Sweden
+46 26-54 20 50
francisco.fraile@itvarena.com

Dan Kopparhed
Interactive TV Arena / KTH
Gävle, Sweden
+46 26-54 20 50
dan.kopparhed@itvarena.com

ABSTRACT
This paper shows the result of three studies at a research site in Sweden. The aim of the studies was to evaluate interactive TV services, using a user-centred approach that includes active validation of usability and accessibility through user testing for development of technology, design and user experience. This study shows the need of early user involvement to decide development areas that have to be in focus. The result also shows that people with functional difficulties are an asset in the development process.

Categories and Subject Descriptors
Interactive TV, user centred development, design for all

General Terms

Design, Human Factors, Experimentation

Keywords
User centred design, accessibility, usability

1. INTRODUCTION
Usability and accessibility in today's interactive TV services have not a high status. One of the greatest influencing factors for this is the distance between users and suppliers of services. In the Nordic countries (Sweden, Denmark, Norway) a small number of companies have virtually a monopoly on the delivery of television services to the consumer.

In the premium television platforms, the set top box (STB) is not selected by the users' preferences, but rather by the service providers' choice with a limited selection of boxes, which lead to boxes with limited use qualities. Developers of hardware and software do not have a direct relationship with end users in the value chain. Instead, they provide hardware and software products to the service provider. Thus, the motivation to engage end users in the design process is rather low.

However, usability is a critical factor in the success of any interactive TV service. Users are very familiar with the traditional use of TV and to integrate new services with the well-known TV user interface is not a trivial task. Initiatives like OpenChoice.tv [1] or Open IPTV Forum [2] aim at opening up the wall around interactive TV services, through the definition of open interfaces for interactive TV services. In this open framework, service providers need to engage users in the design process in order to remain competitive.

2. BACKGROUND
The test and development facilities used in this study consist of a development lab, a usability lab and a Living lab, providing with an optimal environment to engage final users in the design of new interactive TV services. The focus is on IP and web technologies that enable the delivery of content to connected TVs and STBs.

The development lab comprises all the different development tools (middleware distributions, configuration servers, development kits, etc.) needed to develop the services.

The usability lab recreates a living room and is connected to the servers so that it is possible to conduct end user trials in laboratory conditions.

Finally, the services are delivered to a test population over a managed IP platform. The test facilities also include a Digital Terrestrial TV transmitter.

This allows recreating service delivery over two different kinds of platforms: IPTV and Hybrid Broadcast Broadband (HBB) platforms. The main differences between these two platforms is that in IPTV, the interactive services are included in the offer of the service provider and the content distribution is managed, providing Quality of Service, whereas in HBB, interactive services are offered by any internet service provider in the form of TV widgets and content delivery is Over-The-Top, with no guarantees from the IP distribution network. The services under study in this paper regarded an IPTV distribution platform.

3. DESIGN AND STUDY STRATEGY
Our main design and evaluation strategy are based in the ISO 13407[3] human centred design strategy. We have to define what we mean by "Easy to use". Is it just easy to navigate physically or is it easy to understand and easy to read? A simple and effective navigation system in interactive applications is one of the most important things when making it usable and accessible.

We have chosen to involve the target users in the design and evaluation process from the very beginning inspired from User-Centred-Design (UCD) and in co-operative [4] design processes. Very often the end user is not taking actively part as participant in the very design process in UCD projects, but the knowledge of the user is guiding the design. On the other hand, in Co-operative

design processes the end user is participating in the process as in. the project “KidStory” [5] were kids were equal partners in the workgroups. The objective of these workgroups was to build low-tech prototypes with an obvious focus on usability issues close to the participants.

One of our critical objections is towards the use of “common user” or “representative user”, which is one of the issues in this paper and also the thinking that usability testing have to involve large groups of representative users in order to be valid [6]. One solution is to hold down the amount of individuals participating in the evaluation as in Nielsen [7], suggesting the use of as few as 8 - 10 experts in a “discount” usability test to get hold of up to 80% of the usability problems.

Most of the design approaches in order to make easy to use products for all people is to choose representative participants [8][7][9] (common users) to provide ample input to the design process. The group “common users” consists of 'all users'. 'All users' include people with all possible variations of states and conditions [10][11]. There are very few efforts to explore alternative ways of choosing participants in design processes, despite the fact that individual’s functioning is something that can change over time [12] or are tied to a certain situation.

Ohlsson, Persson & Östlin [13] advocate’s that individuals with limitations in their functioning could be a valuable asset in the design process to indicate difficult areas of the usage. To use the concept 'functional difficulties' [14] instead of disabilities might give us another view on the design of products and services for all people that not always have to result in special solutions for certain kinds of groups.

4. METHODOLOGY

This study was conducted as a case study with a triangulation of three separate cases. In case one and two we used an approach consisting a two step testing method. The first step was a “mental model test” (M-test) and the second one was a “usability and accessibility test”(U-test).

In case three we only used the first step; the M-test. All the tests were done in a usability lab.

4.1 Mental model test (M-test)

To evaluate the design from a cognitive perspective we use Donald Norman’s [15] [16] thoughts, were he describes the term affordances to those action possibilities, which are readily perceivable by an actor. This becomes, in our view, a comparison between the user and the developer’s mental map around the design and how the service is thought to be used. In this user test, the participant describes how the layout is perceived to be used and what is expected to happened. This makes the results dependent not only on the physical capabilities of the participant, but also on their goals, plans, values, beliefs and past experience[16]. The participants were only allowed to see the interface and not to manipulate it in this M-test. The participant were asked to reflected on what this interface represent and what were expected to happen in different parts of the interface and how their expected events could be activated.

4.2 Usability and accessibility test (U-test)

The method’s main purpose was to let a user describe how to do a specific task at the same time as they actually performed the task. The procedure allows the participants to express themselves as freely as possible and to let them talk as much as possible from their own point of view during the activity. The test method is based on “Think aloud” protocols [17] and the result is analysed [18] in conjunction with the usability test (effectiveness, efficiency, and satisfaction) [19].

A simple prototype of the interactive applications was used. The usability and accessibility area of this study was primarily targeting the “Ease of understanding”. Only simple navigation in the interactive applications was possible due to the design stage, where not all information was available during testing. In one and two there were two targets of which the activities should reach.

In the first case, after the individual part the participants were divided into two groups where a small workshop was conducted. In this workshop they first discussed the application and then they did a simple paper prototype as a suggestion of improvements.

4.3 The analyses

Both the M-test and the U-test were recorded with video and transcribed and thereafter categorised [20] by described functional arias. The U-test was processed in two ways; first the “Think aloud” was categorised and the second was analysed from the participants’ activity shown on the video. The transcriptions in both tests were done according to Linell’s second level of transcriptions [21].

In case one and two there was also a comparison between their description in the M-test and the U-test in their “think aloud” during their activity. And finally this was compared with their activity shown on the video. In case two there was also a design activity to make a simple prototype [5][4] of their liking, which started with a focus group session. Both these activities were also recorded with a sound recorder.

4.4 Participants

The study was conducted with a total of 35 different participants. In this group there were 20 individuals with defined difficulties; In this case, 11 individuals with mild to moderate intellectual impairments, and 2 individual with eyesight problems, and 5 had difficulties where the hands were affected by motoric impairments; and 2 had difficulties with the Swedish language as their main difficulties.

Eleven individuals were recruited from a school for students with intellectual disabilities (minor). There ages were between eighteen and twenty-one. Among the other participant nine individuals regarded themselves to have some form of functional difficulties. Six of the participants were above 63 years old. The rest did not state that they had any difficulties that affected their use of ITV application.

In case one there were 15 individuals participating; and in case two there were 19 persons participating in the M-test and the U-test and in case three 7 persons participated in the M-test. Only 9 participants considered themselves as technical novices.

All participants were recruited by referrals [22]. The participants were informed that the participation of this study was strictly voluntary and that they could terminate the participation whenever they choose. They were also informed how the material should be used. The information allowed the participants to decide by themselves to participate or not.[23]

5. RESULT AND DISCUSSION

The results are discussed for each case followed by a summary discussion.

5.1 The cases

An interactive television (ITV) application intended to bring easy and understandable information to the guests of municipality driven elderly care centre and the guests' next of kin. This ITV-application should work both on a computer and on STB for TV (Figure 1).

In case one; all participants explained in both the M and the U test "how to navigate" by pointing at areas on the interface, but this differed, unfortunately from the designer's idea of how to navigate. Five of the participants with intellectual difficulties (six in total) in the M-test pointed out the areas of navigation as the major problem area, but not the other eight that participated in this study. In the U-test ten participants of the total fourteen, did have major problem with the navigation. One of the main problems was with the use of the hand held control. The understanding of the navigation keys up-down and right-left and how they affected the upper and lower menu was not clear. This was mainly a problem of different experience of use of this kind of navigation together with a TV. Those who had a setup box at home had also a hard drive, and accordingly were figuring out the navigation faster in the U-test then the rest. All the participants were familiar with hand held controls for the TV.

Figure 1 Screen from elderly care centre

In the M-test a majority of the participants had large problem of describing how the navigation should be working. A third of the participants did not have any clue of how. This is quite unique of this case because the participants usually tries very hard to describe how in this kind of M-test.

One of the main problems in this case was the model of navigation which revealed large differences between the actual possibility to navigate and the participants' model of navigation; and also their actual problem in the U-test with the navigation.

In case two the start menu system (Figure 2) for a ITV-supplier was targeted. The main purpose for the application is to be the start menu in a residence area DTB service. In the M-test the participants described what they saw on the screen and what they expected to happen when interaction with different items of the screen. In the U-test they had two tasks; one to look up some news and the other was to return to the main screen.

Figure 2 Screen shoot from start menu in a residence aria DTB

The result of the differences between the M-test and then U-test was that in the M-test a large majority describing the navigation as it is in a mental model of computer application. *"When I click at the ..... then the next screen is showing....."*. A little more than half of the group did use the arrows and OK button to navigate but the rest tried different approaches before the reach the next screen. All expect two participants had large difficulties to return to the main screen.

The understanding of the menu items was different among the participants. Their understanding changed during their trial and error in the U-test activity. One of the main discriminating factors among the participants was their past experience with similar services, as the users that had experiences of using a STB with a hard drive came faster to an understanding of how the menu system was working.

The main problem we found in this case was the problem in understanding the use of the hand control together with the menu system in the test. It is hard to say in what way this is happening, but as one of the participants said "*I am used to use the mouse to click, but here it is difficult because there is no mouse".*

In case three we only did the M test. We started with a still picture from a talk show. This still picture was then changed and a small symbol was shown in the lower right hand side. This was done three times with three different small symbols, an RSS symbol, a chat symbol (showed as a serial strip text cloud), and a letter symbol.

Figure 3 Screen shoot with RSS svmbol

In the first part with the picture with RSS symbol, Figure 3, we noticed that only two individuals recognised the RSS symbol and thought that this was some form of new feed at the bottom of the screen. They also thought that it was connected to the talk show that was on the TV screen. Some of the participants thought that the RSS symbol was a sign for sound. The rest, 4 individuals, did not have any idea what the symbol could mean or what was expected to happen.

When showing the chat symbol there were four individuals that did not have a clue what this could mean. Two of the participants thought that the symbol was related to comments from the talk show on the TV. And one thought that this was a "normal chat" such as MS-live. One of the main problems with a chat function is the absence of connection to the model of thinking in the computer and in the use on a TV is different as from a user perspective were comments from three of the participants in the test.

The last screen shoot was with a mail symbol in the right-lower section of the picture. Only two participants did the connection of this symbol with a mail function and both participants did the connection that it was meant to be a possibility to send messages to the program. Four participants did an outspoken logical connection of the symbol to a letter, but not to the functionality of mail. Three of the participants also connected the red dot to be a symbol of turning off the TV.

5.2 Summary discussion and future research

When summarizing the result of the M-test it was also noticeable that the individuals with some form of difficulties articulated all

the arguments, except for a couple of arguments that were based on two participant's (both over 63) prior experiences. During the U-test all the use differences were also represented in the group of participants with difficulties. This could have its origin in that this group is more diversified depending on the fact that they have different difficulties. It seems like it should be possible to only have people with difficulties in the test sample. The problematic part is how to define which kind of difficulties that should be recruited to the test sample.

The gap between how the test participants and the developers idea of using the application must be considered in future design processes. It could be a design change, but it could also be something of knowledge transfer that has to be carried out to the user. The mental map gap is a major problem in the accessibility of the functionality

The overall conclusion of these three cases is that there is a need to include the user early in the development process of new functions and graphical interfaces.

To develop innovative server applications can not only focus on what is possible to do with server technology but to have input from the interface design as well as the future user and other stakeholders in this field. We suggest that in future development in the interactive TV area, both at the server side as well as the client side there must be user involvement in the process. Future research is suggested around the user centred process where as many as possible of the different stakeholders take part. The input from the user as well as edge technology should be a part of the process.

6. REFERENCES

[1] OASIS. (2008, Apr.) OASIS Standard. [Online]. http://docs.oasis-open.org/wsrp/v2/wsrp-2.0-spec-os-01.html

[2] Secretariat, Open IPTV Forum e.V. (2009) OPEN ITV Forum. [Online]. http://www.openiptvforum.org/

[3] ISO 13407, *ISO 13407 Human-centred design processes for interactive systems*. Geneva: International Standard Organisation, 1999.

[4] S. Bødker, P. Ehn, D. Sjögren, and Y. Sundblad, *Co-operative Design — perspectives on 20 years with 'the Scandinavian IT Design Model'*. Stockholm: Invited paper, NordiCHI 2000 conference, pp.1-10., 2000.

[5] G. Taxén, A. Druin, C. Fast, and M. Kjellin, "KidStory: a technology design partnership with children," *Behaviour and Information Technology*, pp. 119-125, 2001.

[6] H. Persson and K. Ohlsson, "Unexploited resources in interaction design for Universal Access: People with Impairments as a Resource for Interaction Designers.," in *HCII 2009 universal access*, San Diego, 2009, pp. 145-153.

[7] J. Nielsen, *Guerrilla HCI: using discount usability engineering to penetrate the intimidation barrier*. Orlando, FL, USA: Academic Press, Inc, 1994.

[8] N. Bevan, "Measuring usability as quality of use," in *Software Quality Journal volum 4 number 2*. Springer Netherlands, 1995, pp. 115-130.

[9] J. Nielsen, *Usability Engineering*. London: Academic press limited, 1993.

[10] K. Ohlsson, H. Persson, and O. Östlin, "The concept of Normality and its impact on design of Information and Communication Technology.," in *Proceedings of the biannual Human Factors Network conference in Linköping*, Linköping, 2006.

[11] H. Persson, *Persons with functional diffi culties as resources in ICT design processes, Licentiate Thesis*. Stockholm: KTH, 2008.

[12] P. Gregor and A. Newell, "Designing for dynamic diversity: making accessible interfaces for older people," in *Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing: providing for the elderly* , New York, 2001, pp. 90-92.

[13] K. Ohlsson, H. Persson, and O. Östlin, "The Bollnäs' model for testing, evaluation and design of information and communication technology services," in *Online Deliberation 2005 / DIAC-2005*, Stanford, 2005.

[14] WHO. (2007) World Health Organisation. [Online]. www.who.int/classifications/icf/en

[15] D. Norman, "Affordances, Conventions and Design," *Interactions*, pp. 38-43, May 1999.

[16] D. Norman, *The Design of Everyday Things*. MIT Press: London, 2000.

[17] K. Ericsson and H. Simon, "How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking.," *Mind, Culture, & Activity, 5(3)*, pp. 178-186, 1998.

[18] C. Lewis, *Using the thinking-aloud method in cognitive interface design, Report RC 9265*. Yorktown Heights, NY,: IBM Research, 1982.

[19] ISO 9241-11, *Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability*. Geneva: International Standard Organisation, 1998.

[20] K. A. Ericsson and H. A. Simon, *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press, 1984.

[21] V. Wibeck, *Fokusgrupper*. Lund: Studentliteratur (In Swedish), 2000.

[22] D. Morgan, *The Focus Group Guidebook. Book 1. The Focus Group Kit*. Thousand Oaks, Calif.: Sage, 1998.

[23] Swedish Research Council, Eriksson, S;. (2008, Jan.) Codex, rules & guidlines for research. [Online]. http://www.codex.vr.se/codex_eng/codex/index.html

Current Issues in Broadcasting from a Market Perspective

Artur Lugmayr
EMMi Lab., www.tut.fi/emmi
POB. 541, Korkeakoulunkatu 8
FIN-33101 Tampere, FINLAND
Tel. +358 40 821 0558
lartur@acm.org

ABSTRACT
The world of TV and video undergoes tremendous changes. The medium TV is faced by the challenge of lower advertising revenues, and consumers spending more time on the Internet rather than in front of TV screens. However, also the medium 'Internet' still is experimenting with new revenue streams and many services as e.g. Facebook do not generate revenue. Within the scope of this paper, current trends in the world of TV and video are discussed from a media centered viewpoint. Selected trends in advertising, changing consumer behavior, media theory, and audience are presented and compiled in a SWOT analysis.

Categories and Subject Descriptors
H.4 [**Information Systems Applications**]: General

General Terms
Management, Economics.

Keywords
Television, media economics, digital television, media evolution, marketing

1. INTRODUCTION
TV as media dependent industry relies on heavy infrastructure such as distribution channels, end-consumer devices, content production tools, and initial funding to guarantee operation. Not the mention the high entry barriers considering licensing and coping with regulation authorities. In comparison, the Internet is less regulated, with lower entry barriers, and services can be deployed without relying on a light weight and nowadays well researched infrastructure. But low entry barriers don't essentially make broadcasting the future model for A/V distribution. Within the scope of this paper a few aspects of current issues in broadcasting are discussed resulting into a state-of-the art SWOT analysis of threats and opportunities. This paper is based on an evaluation of articles in trade magazines which acted as related work for performing the study.

On general level, media evolved during the recent decades many times, and to understand the development, it's important to understand the changes in various environments. Within the scope of this article, current trends in TV and video are examined based on the analysis of articles in trade magazines. The conclusion section of the paper discusses these trends, and provides a SWOT analysis of threats and opportunities for the broadcast and video industry.

1.1 Approach and method
The approach of this opened structured study was based on the evaluation of articles published in trade magazines. The study has been based on an inductive approach, where hypothesis emerged during the process of article collection. Articles have been categorized in the categories: *technical*, *content*, *audience*, and *business*. The collected articles were structured as narration and are cited within the scope of this text.

2. McLuhan's "The Media is the Message"
However, both worlds (the Internet and broadcasting) have one thing in common both worlds are evolving according McLuhan's laws of media [14]: extension, closure, reversal, retrieval, and fragmentation. Both media have in common, that they extend some organs of the user (extension) with their audio-visual sensory stimuli.

But there is some downdraft of McLuhan's laws which is a very unwanted effect and currently hits broadcasters hard – especially considering the law of 'Closure'. The Internet takes the attention of the audiences' away from watching TV shows with advertisement breaks. And advertising is still one of the major revenue models in TV. With decreased audiences, broadcasters can't "sell" the audience to advertisers. The challenge for broadcasters is to cope with this trend and cope with the fact that they have to provide services suitable for attracting this type of audiences – especially the one with the age group up to 30 years that spends more time on Facebook e.g.

McLuhan's law of reversal saying that media if developed towards their limit of potential are reversing their characteristics allows interesting developments especially considering the future of TV. Nowadays more and more telecom operators become content distributors with their IPTV channels, and established distributors as cable TV operators are becoming more competition.

Good news for broadcasters is that any new medium retrieves its content from an older media form. Thus the Internet content is just retrieval medium for content before produced for TV. This effect is to be seen e.g. with YouTube, where the new media form relies on an older more traditional media form – video. McLuhan's law of retrieval can be seen as a positive

development, and provides new possibilities for broadcasters to provide their content to audiences.

Broadcasting fragmented rather highly during the previous decades into various sub-forms and genres. There are many good examples for this law of McLuhan. BigBrother or documentary channels are just two of these. Both are subsidiary niches of the general media TV, and successfully attract specific target audiences.

3. CONTENT MODEL AND CONSUMERS

On general level, media evolved from mass media, towards niche media, to participatory media (see [12]). This tendency clearly shows that the consumer moves into the role of being program director and active content contributor. This trend is today clearly visible – today's young generation enjoys their TV shows via web-portals whenever and on whatever device they like; traditional program scheduling looses ground in favor of VoD online libraries; and traditional web service provider offer more and more video online content (e.g. Apple).

Participatory media mean that the consumer moves into the role of being anticipated into the process of compiling content on his behalf. One excellent example to cope with this challenge is the IST LIVE project [4]. One of the project aims was to involve consumers in the creation process of live sport events. A professional journalistic team controlled the process of integrating consumer contributed content fitting to live broadcast shows. This project nicely reflect the grassroot movement that currently takes place in the media environment. The professionals are still in charge of the media editing process; however they keep tightly their consumers involved and let them participate in the creation process.

4. TECHNOLOGY PERSPECTIVE

In these days, TV undergoes many technological challenges after the digitalization of TV as such. One of the trends is to switch from *Standard Definition (SD)* to *High Definition (HD)*. HD offers higher image quality and better resolutions. Currently there can be many examples for HD offerings found, as e.g. aonTV offered by the Telkom Austria providing HD VoD and HD broadcasts [2].

Many governmental media policies developed action plans towards boosting broadband Internet. The European media policy developed the i2010 EU initiative, which already increased the number of broadband households to more than 49% in the EU 27 [3]. A rapid increase in available broadband connection allows the emergence of new players in the world of IPTV and Internet based TV.

Nowadays more 3D technology is emerging on markets and movie block-busters as Avatar lead the way in terms of quality of content and possibilities. There are a few home devices available for 3D content consumption, however, they are in their first generation and lack on providing the essential viewing quality. This is the 3rd 3D wave that is taking place in the world of cinema. However, today's 3D wave differs radically in terms of quality as well as in terms of industrial support to bring the 3D experience to consumer homes. This reflects especially in current standardization efforts and associations dealing with pushing 3D into the homes of consumers. One example is the *Society for Motion Picture and Television Engineers (SMPTE)* defining standards for 3D cinema within the DC28 Committee under the umbrella of the *Digital Cinema Initiative (DCI)* [11]. It's European counterpart, the *European Digital Cinema Forum (EDCF)* examines these standards to cover European requirements for the deployment for digital cinema.

To cope with the task to bring 3D into consumer homes, SMPTE formed a task group examining the challenge to distribute 3D to consumer homes [23]. Other standardization groups such as the *Motion Picture Expert Group (MPEG)* [15] set as their objective to develop standards dealing with the compression, encoding, and transmission of 3D A/V streams. This resulted into a wide set of standards, such as the MPEG-2 *Multi-View-Profile (MVP)*, Stereoscopic MAF ISO/IEC 23000-11 for handhelds (MPEG-4 Part 2 Video), MPEG-4 part 10 AVC (ISO/IEC 14496-10/5e Amd.1), and MPEG-4 AVC Stereo High Profile (ISO/IEC 14496-10/5e Amd.1), 2D+Depth (MPEG-C, Part 3, ISO/IEC 23002-3). Other standardization bodies involved in the development of 3D related standards are *the World Wide Web Consortium (W3C)*, DVB, ETSI, ATIS, Open IP TV Forum, BDA, VESA, HDMI, DLNA, ICDM, ATSC, CableLabs, Open IP TV Forum, etc. Industrial associations such as 3D@Home [9] and the *Consumer Electronics Association (CEA)* [10] attempt to speed up the commercialization process and promotion of technology related to 3D.

5. AUDIENCE BEHAVIOUR

Today, TV is still one of the main platforms for enjoying media content. In the EU 15, people spend almost 4 hours per week in front of TV [20]. The main threat for broadcaster is that more and more spend time in the Internet, rather than in front of TV. Thus the challenge is in keeping the audience attracted to broadcasts, rather than they spend their time on the Internet or providing possibilities to distribute traditional broadcast content on the Internet. Currently broadcasters cope with this trend with providing interactive services on their web-portals or alternatively search for possibilities in partnering up with Internet companies. One example research project is the GAMECAST [21] project aims at converging TV and gaming under one umbrella and letting the audience actively participate in computer graphic pre-generated TV shows.

Predictions, that advertising agencies give very strong forecasts for increases in advertising spending on the Internet [18] we also have to keep in mind, that especially social media services don't bring any revenues. Facebook with 250 million members, and approx. $300-$500 million advertisement incomes don't make any revenue [13]. The business model "free service plus advertising" seems not to be leading somewhere in these cases. Other services, such as e.g. Paltak, which have a broadcast subscription a-like business model, where consumers subscribe to premium services are running profitable. In the case of Paltalk, 5% of premium users generated tens of millions of revenue and profits (see also [13]). This challenge will also remain for broadcasters providing online distribution of their content. More deeper aspects concerning business models will be discussed in a later section of this paper.

Many believe that the Internet is a direct threat to TV. However, this is not fully correct. The Internet consumes the attention of the consumer. Thus consumers are less likely to watch TV and with it watch commercials, still one of the main revenue stream of TV. Thus, broadcasters are forced to alter their program formats in such a way, that they are able to offer to bind consumers to their services in the way they would like to (e.g. utilizing social media). A predominant example is the program format BigBrother by Endemol, which utilizes any distribution channel for attracting and binding consumers.

In contradiction to many opinions, consumers are prepared to pay for professional created online content. This is reflected in a study performed by the Nielson Company published in [8], where only 28% of people are interested in paying for social communities or 24% for consumer generated video. In contradiction, over 50% of consumers where willed to pay for professionally produced video, games, theatrical movies, or music.

6. BUSINESS CHALLENGES

One of the main challenge for TV nowadays, is to provide solutions for providing broadcasting content online. To cope with this, broadcasters have the possibility to provide simple download services e.g. via website portals or hosts (e.g. BBC's TV series, YLE's old news broadcasts). As streaming solutions, broadcasters can provide Web-TV, IPTV, or Peer-to-Peer TV. One of the main considerations in distribution of content via online portals is the associated advertising model, where currently just a few trends are visible. MTV investigated possibilities within the project Inform resulting into different solutions for pre/mid/post roll, overlays, pay-per-view, and classical advertising. Currently consumers mostly accept the first of these models [17].

6.1 Advertising for online video

The business challenges of TV and video are rather tough in the age of Web 3.0. The simple advertising model "attracting audiences, and selling these to advertisers" is currently jeopardized by increasing Internet consumption. Broadcasters are faced with declining audiences. On the other hand, online video (or more generally) publishers also require funding models based on advertising on the same principle. In advertising currently three major business configurations are available for video publishers [22]:

- publisher centric ad-network platforms,
- demand side platforms, and
- ad-exchange platforms.

The current model of "selling audiences to advertisers" is realized via ad-networks and agencies. For marketing purposes ad-networks are hired for marketing campaigns. Ad-networks possess a portfolio of sites, to which they can publish their advertisements. In turn, publishers have the possibility to monetize their size through advertising incomes. One example is the Travel Ad Network [5]. Thus the advertising revenue stream is shared and pricing is controlled mostly by publishers and ad-networks.

But the problems around advertising go even one step further: not only is the competition for audiences a threat for online video publishing, new startups offer solutions, which shift the power for control of pricing for advertising to demand-side platforms (see [22] for a discussion about this problematic). Examples are demand side platforms, where the consumer is able to select his preferred advertisements before watching online video (e.g. Publicis' vivaKi [6] and Havas' Adnetik [1]). This jeopardizes current pricing models completely, as video or broadcast providers have less control over their incomes through advertising.

Ad-exchange based platforms as e.g. Google or Microsoft attract web audiences in large scale, and sell the audience to advertising agencies [22]. Advertisements are placed based on smart algorithms fitting to search results or based on smart algorithm analyzing user actions. In the case of online video a more granular targeted marketing is possible and still in control of publishers.

6.2 Towards media eco-systems

The discussion if either 'content is king' [7] or 'communication is king' [16] is fully obsolete in today's times. Surely, content is the driving force to attract consumers, but in times of convergence and the emergence of multiple platforms for content distribution and consumption we can speak of an ecosystem consisting of devices, content, and communication. Thus data is streamed to media ecosystems, where a single stream is rather an agglomeration of services, A/V data, content descriptions, and interactive elements. A good first example for this development is research work related to the integration of consumer devices.

6.3 Changing value-chain

Today's value-chain for TV and video services is rather complex, and involves many more players than content creators and content distributors. In principle, we can speak of two rivalry junks of chains – the traditional A/V value-chain, and the new Internet value-chain. The main players in the traditional A/V value-chain were creators, distributors, consumers, and aggregators.

In contradiction, the main players in the new Internet based value-chain, many new players were emerging: internet portals, search providers, network providers, mobile service providers, etc. As far both value-chains were separated, however, in recent years the boundaries between were blurred. One example for this trend is Apple TV's initiative to cooperate with CBS Corp. and Walt Disney on offering TV subscriptions via the Internet to compete with cable and satellite TV [19].

7. CONCLUSIONS

The analyzed article envisioned a few trends for the future of TV and video environments. These are compiled in the following section:

- consumers are prepared to pay for professionally produced video;
- Internet services as threat or/and opportunity for broadcasters in attracting more consumers;

- High-Definition (HD) as the next major challenge for distributing TV content;
- rapid rise in broadband connections in consumer homes enable IPTV and web-based TV and video applications;
- fragmentation of the value-chain and specialization of value-chain partners;
- standardization, commercialization, and promotion of bringing 3D to consumer homes;
- matching of advertising models for online A/V distribution matching current and future B2B standards;

The basic results of the study have been compiled in Table 1 within a SWOT analysis or strength, weaknesses, opportunities, and threats for broadcasters.

Table 1. SWOT analysis of current trends in broadcasting

STRENGHT	**WEAKNESS**
High entry barriers and capital investment Additional revenue streams through multi-channel and multi-format distribution TV is still a well accepted mass medium Well established advertising models New distribution channels allow content re-engineering and enable new revenue streams	Non-sufficient payment offers Lack of program formats coping with participatory consumers Strong regulatory framework and diverse governmental policies Internet takes attention of consumers away from TV Rivalry between advertising platforms on the Internet Lack of program formats suitable for multi platform and interactivity
OPPORUNITIES	**THREATS**
Internet as additional distribution channel Content diversification through new technologies Premium content offerings (e.g. HD, 3D) Broadband as additional competition to existing TV distributors Consumers are prepared to pay for online Merger and cooperation between both, traditional and new value chain partners Fragmentation of value chain easies entry barriers	Internet with service offerings substituting TV services Advertising budgets are spent on the medium Internet Young audiences prefer other media Lack of adequate advertisement platforms Demand side advertising platforms determining pricing for advertising Participatory content models are not adaptable in a TV environment 'Internet services are free' mentality difficult to adapt for established media companies

REFERENCES

[1] Adnetik. http://www.adnetik.com.

[2] aontv. http://www.telekom.at.

[3] Eurostat. http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/.

[4] Ist live. http://www.ist-live.org/.

[5] Travel ad network. http://www.traveladnetwork.com/.

[6] Vivaki. http://www.vivaki.com/.

[7] King content. *The Economist*, 2006.

[8] By the numbers. *Mediaweek*, 20(4):42, January 2010.

[9] 3D@Home.

[10] C. E. A. (CES). www.ce.org.

[11] D. C. I. (DCI). http://www.dcimovies.com/.

[12] L. Kueng. *Management in the media - from theory to practice*. Sage, 2008.

[13] D. Lyons. It's time to pony up. *Newsweek*, page 17, 3rd August 2009.

[14] M. McLuhan, H. M. McLuhan, and W. T. Gordon. Gingko Press, 2004.

[15] M. P. E. G. (MPEG). http://www.chiariglione.org/mpeg/.

[16] A. Odlyzko. Content is not king. *http://www.research.att.com/ amo*, 2001.

[17] O. Otlacan. http://www.adoperationsonline.com/2009/07/17/shorter-pre-roll-plus-lower-13-tops-among-best-ad-units-for-short-form-online-video-according-to-new-mtv-networks-study/.

[18] A. Research. Online video advertising spending in asia, europe and the us 2008 and 2012. *Financial Times*, July 15th 2007.

[19] Y. I. K. Sam Schechner. Apple tv-service proposal gets some nibbles. *The Wall Street Journal*, 22nd December, 2009.

[20] SCB. Harmonised european time use survey - secondary activities. https://www.testh2.scb.se/tus/tus/StatMeanSact.html.

[21] T. Schmieder, R. J. Wierzbicki, and A. Lugmayr. Gamecast - a cross-media game and entertainment system. In *Changing Television Environments, TICSP Adjunct Proceedings of EuroITV 2008*, volume 42, pages –, Salzburg, Austria, 2008. Tampere University of Technology (TUT).

[22] M. Shields. Losing control? web publishers fear agencies' demand-side platforms. *Mediaweek*, 19(42), November 2009.

[23] SMPTE. Report of smpte task force on 3d to the home.

Audience measurement system for digital television

Valdecir Becker
Laboratório de Sistemas Integráveis
Departamento de Engenharia de Sistemas Eletrônicos
Escola Politécnica da USP
Av. Prof. Luciano Gualberto Trav. 3 n. 158
05508-900 SP São Paulo Brasil
valdecirbecker@gmail.com

Marcelo Knörich Zuffo
Laboratório de Sistemas Integráveis
Departamento de Engenharia de Sistemas Eletrônicos
Escola Politécnica da USP
Av. Prof. Luciano Gualberto Trav. 3 n. 158
05508-900 SP São Paulo Brasil
mkzuffo@lsi.usp.br

ABSTRACT
The introduction of digital TV results in problems for audience measurement. New methods need to be developed for determining the audience in multicast environment in which there exist various digital channels in the same frequency. These new methods can be based on software systems, which permit new measuring techniques that can improve the measurement quality. This paper presents an audience measurement system for digital television, which is based on software and real time data collection.

Categories and Subject Descriptors
H.4.0 [**Information Systems**] Information systems applications – General.

General Terms
Human Factors

Keywords
Digital TV, Audience Measurement.

1. INTRODUCTION
The introduction of digital TV results in problems for the audience measurement. Transmissions in multicast environment, with various channels in the same frequency including mobile transmission, demand new audience measurement methods. These new methods can be based on software systems, which allow more efficient data collection for statistical analytics. Such systems permit new measuring techniques that improve the measurement quality.

Usually, the digital television introduction studies focus on the consequences in areas like image quality, interactivity, usability and new business models. Academic research studies about digital TV audience are quite rare. In countries where digital television is only standard or high definition, with a little interactivity, and without multicasting, the same measurement techniques can be used.

However, if the broadcaster opts to use more standard-definition (SDTV) channels instead of one high definition (HDTV) signal, new techniques are necessary to measure the audience of these new channels. The analogical measurement is unable to identify the separate channels inside the same frequency.

This paper proposes and presents a new audience measurement system, capable to get information in digital TV environment including multicasting channels and, in the future, interactivity use.

2. TELEVISION AUDIENCE MEASUREMENT
During the twentieth century, the increasing participation of private companies announcing their products in radio and television was directly proportional to the development of tools of audience measurement. Audience research has thus born as an instrument of control of the advertising activity [1].

Moreover, in business models based in advertising, where the content is free, the pressure for knowing the number of audience and for developing measurement tools is stronger when the information is scarcer. The increase of the transmission technology brings more consumers for the content.

The television is essentially a mass media, and as such, needs audiences to grow. The audience measurements supply an immediate necessity of the TV companies, which need to legitimize the values used in the negotiations of sponsorship and sales of advertising space. Without this validation, they would also lack parameters for definition of monetary values for different programs or channels.

The knowledge about the audience rates is practically the only valid criterion for the sales of spaces for the advertising agencies and advertisers who are the ones that definitively finance the radio and the television. [2]. This is possible thanks to the evolution of the measurement methods that the number of audience has a central role today, as much in the planning of the communication vehicles as in the commercialization of advertising spaces. No system, especially those in which the communication has a vital role, can operate without an audience investigation or another good feedback tool.

Commercially, television is the communication system with the highest penetration in the Brazilian market, with more than 97% reach [3]. It is the preferred medium by market professionals in Latin America, concentrating almost 49% of the all advertising expenditure [4]. However, TV is a very expensive medium: a simple 30 seconds advertising insertion can cost R$ 381K (around U$ 220K) in the prime time of the main channel [5].

2.1 Measurement in Brazil
In Brazil, it is the Brazilian Institute of Public Opinion and Statistics – Ibope - that measures the audience [6]. The company uses a peoplemeter, an audimeter connected to the television. It

has integrated a tuner, which recognizes the channel watched by the viewer. The numbers one to eight, arranged in peoplemeter's display, serve to identify which of the occupants of the house is watching television at that time, allowing some analysis of public profile. The identification is made by the user by means of the remote control.

Figure 1. Peoplemeter audimeter, with remote control

The most important measurment market in Brazil is São Paulo, with a statistical sample composed by 700 households, sending real time information, with one message every minute. In total, Brazil has 3,500 audience measuring points, in 11 cities. The data is sent to Ibope through radio frequencies and is processed by the company's software before the sending to the costumers.

Ibope makes available to subscribers of the hearing service on TV a Software Media Workstation, accessible via the web. The software includes information about the viewers' behavior, which make possible detailed analysis. The program allows the user to define filters for the desired analysis. Thus, it is possible to analyze, for example, by broadcaster or slot, if one wants to compare programs, or by the audience of programs, if one prefers to check the penetration of that particular program or station at a target audience. The target audience is one of the possible filters of the analysis that can be done using profiles based on age, sex and social class. It is possible to set rankings between programs from targeted audience, and a complete study about the program, with data from the every minute audience.

Media Workstation is offered in two versions of software: Media Workstation Standard and Premium Media Workstation. The first allows analysis from 15-minute blocks of these variables: audience, reach, share, total connected, profile and Gross Rating Point (GRP)[1]. The Premium Media Workstation, which allows a minute by minute analysis, includes a wider range of variables.

3. AUDIENCE MEASUREMENT PROBLEM IN DIGITAL TV

Brazilian digital TV transmissions started at December 2007, in São Paulo city, based on Japanese ISDB system. The most important broadcast companies launched digital transmissions, alternating SD and HD programs. On that occasion, the One-Seg transmissions, for portable reception, were also started. Nowadays, 29 cities already have digital signal, which comprehend 60% of the Brazilian territory [7]. The Brazilian television is based on 6 MHz channels, using both Very High Frequency (VHF) and Ultra High Frequency (UHF). The digital transmissions use only UHF channels.

The audience measurement in television is based on the frequency tuned by the TV set. In Brazil, the measurement is done by an audimeter, which identifies the tuned frequency and sends the information to a database center. For example, if the viewer tunes the channel two in analog TV, that is TV Cultura, the peoplemeter get the information about frequency (54 to 60 MHz) and send it to Ibope's database. The time, share and other statistics about audience are calculated from this database.

In digital transmissions, every broadcast company received an additional channel for simulcast transmission. In the Brazilian system, the tuning is done with the same numbers that were used in the analogical one, for easier memory. It eliminates the need to memorize the TV channel numbers anew. It means that TV Cultura continues to be tuned in channel two, despite the transmission in frequency 530 to 636 MHz, which is 24 UHF.

On of the most important differences between digital and analog transmission is the multicast possibility. In analog TV, only one television channel is possible in each 6 MHz frequency. The measurement is made identifying this frequency (Figure 2).

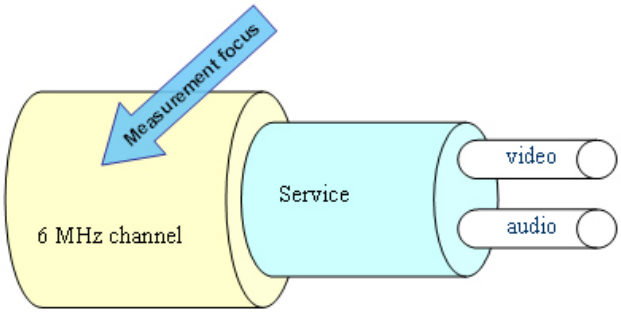


Figure 2. Measurement focus in analog TV

With digital television, it is possible to transmit more than one TV channel in each 6 MHz frequency. Brazilian system, with the H.264 codec option, permits up to eight SDTV channels in that frequency, or one HDTV and two SDTV. These additional digital channels, transmitted together are named logical or virtual channels. The logical channels can not be identified in measuring techniques based on the frequency.

The TV Cultura is the only channel that has multicast on air now in Brazil. It is transmitting two additional channels: Univesp TV and Multicultura, besides the usual channel and the One-Seg, for mobile reception, but able to be tuned with fixed TV sets. The Cultura TV schema is showed in figure 3.

If we apply the frequency based techniques to digital TV, it would be possible only to identify that the channel 24 is tuned. It is impossible to identify what logical channel the viewer is watching. There are three more virtual channels that could be tuned.

[1] GRP is an advertising term, used to measure the size of an audience reached by a specific media vehicle or schedule. It is the sum of audiences reached by a show or channel.

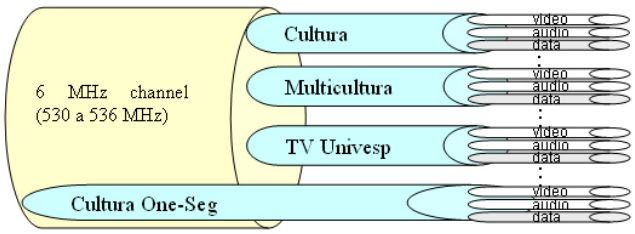


Figure 3. TV Cultura digital transmission schema

3.1 International solutions

European countries, especially Spain and England, and digital paid TV systems, have faced the same problem for measuring the audience. Some different methods have been used.

One of those solutions is signal codification, where the broadcaster inserts a code in video signal that identifies the channel and the program. This code can be an inaudible audio or an invisible noise in video. However, this system has a considerable limitation. The dependence of contribution of all companies, and standardization necessity for emission of these signals make the system difficult to be deployed in large markets.

Another possibility is the video signal recognition. It is currently used by Picture Matching System (PMS), predominant measurement system in Spain [8]. It works with the content of the video signal, that is, with the image being shown on the screen. The process entails storing samples of the video signal that will be compared with samples of the signal emitted by each channel, thus identifying which channel is syntonized.

The audimeter does not compare directly the video signals of the samples, but performs multiple comparisons of brightness between two samples. Therefore, the image is reduced to 32 small parts, whose average brightness is compared. These 32 parts are organized into 16 pairs, from which it can get two possible outcomes, depending on which part is darker. The result is 16 bits of information per screen (1 bit for each pair).

The problems of identifying the broadcast station appear with the interactivity. The system is not able to identify when an interactive application overlays the video, which creates distortions.

In France, companies France Telecom and Mediametre developed a measurement system capable of analyzing both analog and digital signals, called Module Cable Satellite (MCS). However, the detailed operation of the system has not been fully published yet [8]. The MCS identifies radio waves to make a first analysis. If it is an analog signal, the audience is attributed to the corresponding channel. For digital, the system identifies the set of channels transmitted in the stream and then the individual channel which is being tuned.

Besides the systems described above, other software based techniques have been developed that complement or add functions to the audience measurement in digital TV, specially paid digital TV, making up the deficiencies of other techniques. In this sense, they are not exactly audimeters, but systems that use resources of the receptors to identify the tuned channel. The data collection is done by specific software installed for this purpose in the receiver. Information is sent by the return channel and analyzed by software called InfoSysTV [9]. The great advantage of this system is the possibility of measuring a large sample, at low cost. The market research company TNS [10] is using this type of system.

4. NEW MODEL BASED ON SOFTWARE

To suggest another solution for the problem, we developed an audience measuring system based on software. The system identifies what logical channel is tuned and sends the information to a database using a set top box return channel. In future, it will be expanded to read information about the interactivity being used. It is based on Brazilian digital free to air TV, using middleware Ginga.

From the architectural point of view, set top box is a computer. It has memory, processor and storage capacity and it can also have internet access. Thus, the information collected by the software can be stored or sent immediately to a database, with an internet connection.

Figure 4 Welcome and identification screen (translated here from Portuguese to English for illustrating purpose).

The system identifies the logical channel reading the Network Information Table (NIT) from the MPEG-2 system transmission. The NIT carries information about streaming videos of the channel. The NIT conveys "the physical cluster transport streams (TS) existing on the same network and its characteristics, as well as relevant information about the tuning of existing services" [11]. Through this table it is possible to map all the services offered by a particular broadcasting channel and which are being tuned by the receiver.

When switched on, the receiver connects to the database, authenticating the measuring point. After authentication, database sends the profile information of users registered at this point, which are displayed on the TV screen. Then user has to choose his profile, enter information about the number of people accompanying him in audience and press OK. This screen is shown in Figure 4.

After identification, system reads the tuned channel and sends the profile information, number of people and channel to database. Thus, the first record of the audience is done. After that, system stands away for a channel exchange or TV shutdown. When one of these actions occurs, the result is sent to the database, updating the information.

In this version, the information that must be provided is the profile and the number of people in front the TV. But it is possible to collect any information desired to refine the identification and improve the measuring of TV viewing behavior. The available profiles are pre-defined statistically, considering the residents that compose the sample. If it is not a resident turning on the TV, he should select "visitor" in profile.

In database are made all the calculations relating to the conduct of the hearing, as an average audience, syntony average time, tuned channels, syntonization times, connected receivers, audience profile, reach and GRP. The database implementation is similar to the Ibope Standard Media Workstation. User selects the display options through menus. The results are automatically updated every 15 minutes. In this version, the database does not generate the minute by minute results.

Figure 5 shows an audience and share screen, accessed via the Internet. In broadcasters' columns (Emissoras), the top line represents the program average audience of the period, in relation to all recipients composing the sample. Also considered are tuning time of each station and quantity of receivers. This percentage is known as the audience of the program.

The bottom row represents the percentage of the sample receivers tuned to the channel in relation to all connected receivers. This percentage is known as program share. The update is done automatically every 15 minutes, when the first line leaves and the new one enters at table bottom.

In this prototype, the software is placed in the set top box and has no interference in the viewing habit (except the first identification screens, described below) nor is there any need for additional remote controls, as is the case with peoplemeter today. Another option to use this software is to transmit it multiplexed with the digital signal, and use the Ginga middleware to execute it in the set top box. In this case, the broadcast company can define exactly what kind of information is relevant in different programs or timezones.

Sistema de Medição de Audiência Poli/USP									
Horário		Emissoras							
Inicial	Final	04	05	07	09	11	13	15	TOTAL
09:00	09:15	100 100	0 0	0 0	0 0	0 0	0 0	0 0	100 100
09:30	09:45	0 0	100 100	0 0	0 0	0 0	0 0	0 0	100 100
10:00	10:15	60 60	40 40	0 0	0 0	0 0	0 0	0 0	100 100
10:30	10:45	13 13	87 87	0 0	0 0	0 0	0 0	0 0	100 100
11:00	11:15	25 25	75 75	0 0	0 0	0 0	0 0	0 0	100 100
11:30	11:45	0 0	100 100	0 0	0 0	0 0	0 0	0 0	100 100
12:00	12:15	34 34	66 66	0 0	0 0	0 0	0 0	0 0	100 100
12:30	12:45	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
13:00	13:15	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
13:30	13:45	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0

Figure 5. Audience and share screen.

5. CONCLUSIONS AND FUTURE WORKS

The system of audience measurement presented in this paper offers a promising solution for improving current techniques. Compared to the system used by Ibope in Brazil, the framework proposed in this paper is more agile, flexible and cheaper. The fact that it is based on real time data collection, adds velocity offering faster results.

Compared to European measurement systems, this technique dispenses any type of inserts in the audio or video signal, which allows a complete separation between the operations of the measurement company and TV stations. The system proposed in this paper can be replicated and used in any digital TV system that uses multiplexing MPEG2.

5.1 Future works

For further work is necessary to add value to the database, to offer new services, especially information about the behavior of the audience in every minute. This analysis is essential for live shows. Interactivity measurement is another requisite improvement. Finally, it is important to test the system with more set top boxes, sending information simultaneously.

6. REFERENCES

[1] BERMEJO, Fernando. The internet audience: constitution & measurement. Nova Iorque: Peter Lang, 2007.

[2] IGLESIAS, Francisco. Investigación de audiencias en televisión y mejora de calidad de los contenidos. In Documentación de las Ciencias de Za Información, IX. Edit. Ijniv. Compí. Madrid, 1985.

[3] IBOPE, Target Group Index. Banco Ano 7, onda 2 e Ano8, onda 1 (fev/2006 a jan/2007). Penetração de Meios e Classes sociais – Mercado Nacional (principais regiões metropolitanas). Total população: 12 a 64 anos

[4] IBOPE, Monitor Plus. Meios: TV Aberta (29 mercados), Revista, Jornal, Rádio, Outdoor e TV por Assinatura. Banco utilizado: 1ª remessa de jan/07. Valores base em tabela de preços dos veículos. Informações referentes ao período: 01/jan/2006 à 31/dez/2006

[5] JOVE: Tabela de preços de comerciais das emissoras de TV aberta de São Paulo. Available: http://www.jovedata.com.br. São Paulo, 2009. Accessed 01/12/2009.

[6] IBOPE. Instituto Brasileiro de Opinião Pública e Estatística. Available: http://www.ibope.com.br. São Paulo, 2009. Accessed 01/12/2009.

[7] SBTVD. Sistema Brasileiro de TV Digital. Available: http://www. www.forumsbtvd.org.br. São Paulo, 2009. Accessed 01/12/2009.

[8] AIMC. El uso de la tecnología en la medición de las audiencias. Spain, 2008.

[9] BBM Analytics. InfoSysTV. Available: http://www.bbmanalytics.ca/infosys_tv.html. Acessed: 28/11/2009.

[10] TNS. Available: http://www.tnsglobal.com. Acessed: 28/11/2009.

[11] ABNT, Associação Brasileira de Normas Técnicas. ABNT NBR 15602-3: Digital terrestrial television – Video coding, audio coding and multiplexing Part 3: Signal multiplexing systems. Rio de Janeiro: ABNT, 2007.

What do you want to watch (again)? Video Navigation Using Recency Frames

Diogo Pedrosa Didier A. Vega Maria da Graça C. Pimentel
SCC-ICMC-Universidade de São Paulo, São Carlos-SP, Brazil
{diogo,davo,mgp}@icmc.usp.br

ABSTRACT

Inspired by proposals which exploit the episodic memory for channel retrieval, in this paper we present a model that allows navigation between channels (spatial navigation), and the replaying the watched content (temporal navigation). Our approach is based on capturing frames while users watch a particular program, and presenting the captured frames on a timeline designed to show recently watched frames. Our model offers users the possibility to watch again any recently seen content, and the possibility of watching the event through the end, even if it was not completely watched in the first time. We present a proof-of-concept prototype designed and built to run on a digital TV middleware, and discuss the results of a usability evaluation.

Categories and Subject Descriptors

H.5.1 [**Multimedia Information Systems**]: Video; H.5.2 [**User Interfaces**]: Interaction styles, Graphical user interfaces; H.5.4 [**Information Interfaces and Presentation**]: Hypertext / Hypermedia - Navigation

General Terms

Experimentation, Design, Human Factors

Keywords

Video Navigation, Interactive Digital TV, Brazilian Ginga Middleware, NCL, Lua, HCI, Episodic Memory

1. INTRODUCTION

Because Digital Television (DTV) allows a large number of channels, there is a demand for special tools for helping the user to search for a TV program (e.g. [6]) and to support navigation (e.g. [10]).

Episodic memory refers to memory for personal events and the temporal-spatial relations among these events [12]. Cognitive science studies [13] [12] show that the episodic memory is responsible for storing events from the everyday life and that, although it allows automatic event storing instantaneously, the corresponding retrieval is difficult because it is known to be dependent of environmental stimuli. In the context of Digital Television, Kim et al. [7] suggested that the channel and the moment in which the content is watched are examples of information stored in the episodic memory, and they proposed the presentation of thumbnail frames of the video programs the user watches in a timeline built to serve as a cue to retrieve the information of what was seen in which video (channel).

Inspired by the work of Kim et al. [7], in this paper we present a model that allows spatial navigation and temporal navigation. By spatial navigation we mean the navigation between channels, and by temporal navigation we mean the navigation by replaying the watched content. Our approach is based on capturing frames while users watch a particular program, and presenting the captured frames on a timeline designed to show recently watched frames. The model explores a scenario where users have the possibility to review any recently watched content. This scenario also offers the possibility to watch a video through the end of the event, even if it was not completely seen in the first time. Our work extends the proposal presented by Kim et al. [7] both by addressing the problem of the availability of the replayed content, and by further defining rules associated with the inclusion of new frames in the timeline. We evaluated our model by implementing a proof-of-concept prototype that runs on a digital TV middleware which offers an open API for developers [11].

After reviewing relate work in Section 2, we present our model (Section 3) and a proof-of-concept prototype (Section 4) along with its usability evaluation (Section 5). In Section 6 we present our final remarks.

2. RELATED WORKS

In the context of video retrieval in large databases, the need for special tools to aid the retrieval task has lead to several contributions, such as alternatives for browsing among video thumbnails [2]. In the context of video browsing in mobile devices, the timeline-based main navigation has been extended with additional operations demanded by the size of the devices [5].

Crockford and Agius [3] made an empirical investigation to understand how users relate to a VCR-like control set and how useful it really is. With the advent of DVD and personal video recorders, new temporal navigational models came to assist viewers in the search and navigation. The model proposed by Drucker et al. shows thumbnail frames captured at regular intervals and arranged in a timeline. Usability tests reported that the proposal by Drucker et al. was the most pleasant model, despite the poor results in terms of performance [4].

Frames that represent a certain time interval of the video are known as key-frames. Arman et al. [1] have employed key-frames to allow users to find a particular point in a sequence, or to decide whether the sequence content is relevant to their needs.

While previous work investigated issues related only to temporal navigation, Kim et al. [7] addressed difficulties related to navigation both in the space and time dimensions.

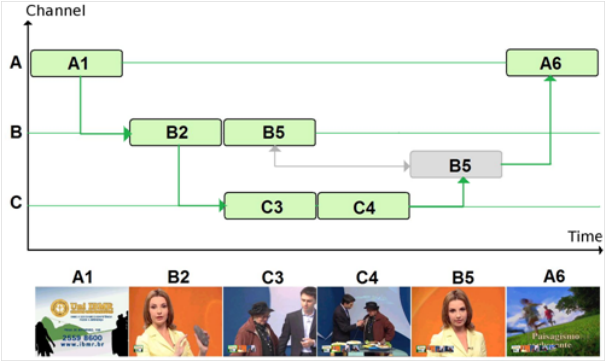


Figure 1: Sequence of frames captured.

Based on episodic indexing theory, they designed and tested two navigation aids: recency-frame and short reminder. The recency-frame consists in presenting a series of small thumbnail images along the time dimension, as other key-frame models do, but it differs from other models by the fact that it specifies the presentation of both frames recently seen on the current channel and on other channels. The short reminder approach proposed by Kim et al. [7] presents to the users, as they return to a channel, a small-screen video consisted of the last 5 seconds of video watched in this channel. The experiments have shown that both aids allowed a better user satisfaction, but the gain provided by the recency-frame aid was significantly higher [7].

3. EXTENDED RECENCY-FRAME MODEL

The extended recency-frame model is detailed in the example illustrated in Figure 1, in which a user interacts with a tool that implements the model. The user starts watching channel A, which causes the frame A1 to be captured. Then, the user switches to channel B, and later, to channel C, adding frames B2 and C3 to the timeline. The user remains in channel C long enough for the tool to capture a new frame in that channel: C4. The user then decides to return to the point where the frame B2 was watched, and indicates that the video has to be played back from frame B2. The tool is executed continuously and capturing frames no matter the source, so frame B5 is captured. It is worth noticing that frame B5 was originally broadcast before frame C4, but it is displayed after frame C4 on the timeline, because the timeline respects the order in which the frames were watched. Finally, the user decides to switch channels and to watch channel A, which causes the capture of frame A6. The model also allows users to watch the broadcast signal of the channel whose content is being played.

Our extended recency-frame model provides five basic functionalities to the user: to open the timeline; to close the timeline; to navigate between saved frames; to select a frame to cause the playback of the video associated from the point where that frame was captured; and to return to live broadcasting. A state diagram covering these functionalities is presented in Figure 2.

In the recency-frame model proposed by Kim et al. [7], the frames added to the timeline link to the last scene presented to the viewer at each channel change: the idea is to allow users to continue watching exactly from the point where they had stopped.

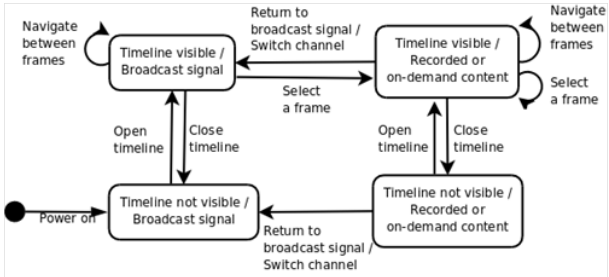


Figure 2: Functionalities in the extended model.

It is very important to observe that our work, extending the recency-frame model, provides marks that facilitate the replay of a desired content. In the model presented in our work, each time a new channel is tuned for at least 3 seconds, a new frame is added to the timeline. This duration derives from the fact that this is the time needed to the information to be stored in the episodic memory [7]. Moreover, the extended recency-frame model also defines a capture rate of a new frame every 7 minutes while the user stays in the same channel. It is worth noticing that these times should be configurable and may vary with the type of program and characteristics of users.

The amount of visible frames at any given moment should be enough to allow the user to visualize frames linked to a reasonable duration of the watched content. This amount is also related to the size of the thumbnail frame needed to allow the scene to be recognizable. The prototype application we present shows at most nine frames at the same time, but we recommend that the application should allow these values to be configured by the users. To allow users to watch again, and until the end, any content partly and recently seen, the application must adopt a strategy that lets the content to be always accessible. One way to do this is to keep recording the events of all the channels in which the user browses. In the worst case, this demands one TV tuner for each available channel, which is likely to be too expensive for most users.

Our strategy is to use video on demand services when the user selects a frame in the timeline, thus, avoiding the need of numerous tuners. In fact, our model suggests to record only the contents actually seen by the viewer, and to use a video on demand service to show the non-recorded part of the content. To be able to use video on demand services, the tool needs to store also metadata from electronic program guide (EPG), such as the name and the current position of the program. This approach also demands a STB with only one tuner, and it has the additional advantage of not requiring video on demand services when the user just want to watch again a content recently watched.

4. PROOF-OF-CONCEPT PROTOTYPE

This work explores the recency-frame model as a resident application prototype in one of the existing Digital TV middleware Ginga, the middleware of the Brazilian Digital TV System. The Ginga-NCL environment is responsible, in Brazilian Digital Television System, for the processing of declarative applications coded using the NCL declara-

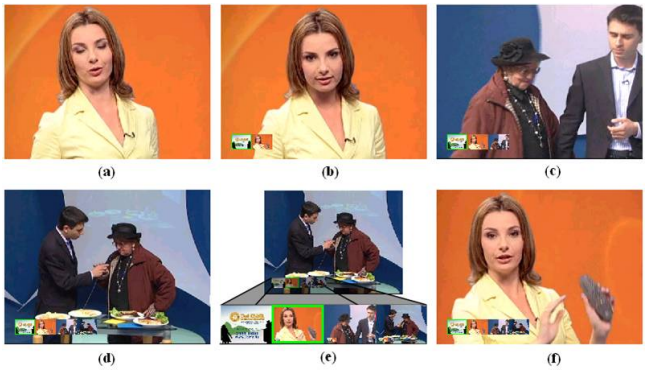


Figure 3: The prototype in use.

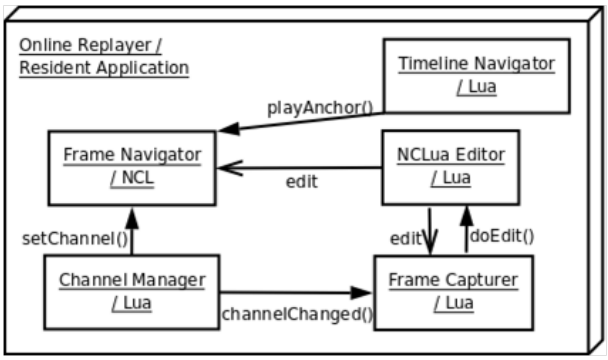


Figure 4: Communication diagram.

tive language[1]. To extend the declarative environment, the Ginga-NCL module adopted the extension language Lua, which offers support to procedural programming in general.

A sequence of screenshots showing the interface of the prototype application is shown in Figure 3. The user starts watching a channel (Fig.3.a) and decides to open the timeline, which has already 2 captured frames (Fig.3.b). With the timeline open, the user decides to change channels, which provokes a new capture after 3 seconds (Fig.3.c). As the user watches the same channel for more than 7 minutes, the tool captures a new frame of the same channel. At this point, the user presses the left and right arrow-buttons to focus a frame associated with the content he wants to watch again (Fig.3.e shows an enlarged view of the timeline, and the second framed highighted as a result of the selection). Finally, the user selects the second frame: the playback of the associated content starts from the moment in time associated with the capture of that frame (Fig.3.f).

The communication between the modules of the application is represented in Figure 4. The application, called *Online Replayer*, has as its entry point the NCL document (*Frame Navigator*), whose function is to control the presentation of media objects, defining the screen area where they are presented and when they are initialized or finished. One of the media objects of the document is a Lua script, which is composed of four modules that are responsible for controlling the channel switching (*Channel Manager*), providing a way to capture frames (*Frame Capturer*), allowing access to recently seen content (*Timeline Navigator*), and editing the application files (*NCLua Editor*).

To simulate video channels in the virtual machine with Ginga-NCL, the *Online Replayer Application* starts in parallel three videos of 45 seconds each. It displays only one video at a time and restarts them when they reach the end. The *Channel Manager Module* uses Lua's *Event Module* to listen to key events and to manage channel changes. At each channel change, the *Channel Manager Module* posts two events to be processed by the NCL document (represented by the *setChannel( )* call in Fig. 4): one event to hide the video and to mute the audio representing the current channel, and the other event to show the video and to unmute the audio representing the new channel.

The *Frame Capturer Module* is initialized by the NCL document when the application starts. Its function is to capture frames. To do that, it is notified by the *Channel Manager Module* at each channel change (*channelChanged( )* call in Fig. 4).

The function of the *Timeline Navigator Module* is to handle keystrokes, allowing the user to change focus between frames and select a frame to watch the linked content (*playAnchor( )* call in Fig. 4).

To be able to order the presentation of the content linked to the selected frame, starting from the point in which the frame was captured, the media node representing the channel must have an anchor indicating this moment. That makes clear the need to edit the NCL document each time a new frame is captured, while it is been exhibited.

To allow NCLua scripts to edit an NCL document, the Ginga-NCL standard defines the *edit event class*. However, the current version of the reference implementation [8] does not support this class of event, and the *NCLua Editor Module* needed to be implemented (as listed in Fig. 4). It uses text files manipulation functions provided by the standard Lua API. When the edit event class is made available by the reference implementation, the *NCLua Editor Module* should be updated.

The problem with this solution is that when a NCL application is started, the Ginga-NCL presentation engine parses its documents and builds a Document Object Model (DOM) tree in memory [10]. Hence, just editing the NCL file would be useless. To work around this problem, the adopted solution restarts the NCL application every time the NCL document is edited. This is the function of the *Presentation Manager Component*; this is a temporary approach – recent updates on the reference implementation tackle this issue.

To increase performance, edits (*doEdit( )* call in Fig. 4) to the NCL document are delayed until they are really needed, that is, when a frame is selected by the user.

5. USABILITY EVALUATION

Given the current stage of our prototype, we carried out a usability evaluation of the prototype using the heuristic evaluation, one of the discount usability evaluation methods proposed by Nielsen and Mölich [9] by three usability specialists.

The specialists have identified seven problems which we list next, along with a discussion regarding each problem:

Lack of navigation history: The tool offers a history of what the user watched and allows him to use it to access some content again. If the user navigates in a set of non represen-

[1]http://www.ncl.org.br/

tative captured frames, he might try to select some frames of the timeline looking for a specific part of the program. Also, a simple indication of whether the frame was already revisited or not might be useful.

Lack of highlight of the current frame in the timeline: The idea of indicating the frames that are related somehow to the video being present is interesting. Frames from the same program or the same channel could be marked. Besides that, the selected frame should be the focused one after the restart of the tool.

Lack of information relative to the size of the video and the progress of its exhibition: This issue is related to the fact that our tool should be integrated to an electronic program guide, as both resident applications offer the possibility to navigate and select content. Normally, the EPG already shows the current position and the duration of the program. Our intention was to propose and validate a complementary way of video navigation, not to implement a complete program guide.

Limited timeline length: This is a strong and known limitation of the prototype, and will be fixed in the next version. The bookmark suggestion is also interesting. But, as the functionality suggested in the preview problem, this one is also offered by EPGs or equivalent resident applications that have content recording capabilities.

Lack of mechanisms to ensure user privacy: The requirement to remove frames from timeline, not elicited in the first design, should be considered. Not just frame by frame, but also options like "all frames from this channel" or "all frames since the television was turned on".

Low contrast between timeline and video: This issue should be addressed in the next version. A better way to indicate when no frame has been captured should also be designed.

Absence of help: Although the tool is very simple, the user needs some explanation about the purpose of it and the functionalities it has. A container to the timeline, as suggested to solve problem 6, can also be used to aggregate some instructional text.

The task performed by the specialists was to evaluate the prototype, and the problems and new requirements identified will be tackled in future versions of the prototype. However, it is important to observe that none of the problems identified by the specialists in the prototype are derived from the Extended Recency-frame model itself – we argue that this result validates the model.

6. FINAL REMARKS

This work discussed a model that allows the user to easily watch again any recently seen content, and offers the possibility to watch it through the end of the event, even if it was not completely seen in the first time. This is important due to the task complexity of finding the desired content caused by the availability of a large number of channels, the availability of video on demand services, and the several alternatives available to record contents.

The feasibility of the model was evaluated by means of a proof-of-concept prototype developed as a NCL/Lua resident application, which runs on a Ginga middleware. The prototype allowed a heuristic evaluation to assess the user-interaction with the application. Besides identifying usability problems, the evaluation produced new important requirements which should be tackled in the next version of the prototype. Overall, the work validates our extended recency-frame model and identifies important points for its evolution.

Regarding future work, we plan to tackle both the problems and new requirements identified by the specialists in the usability evaluation. We also aim at working to increase the performance of the solution used to capture frames to allow a frame capture rate far bigger than the rate of adding frames to the timeline. This would give a range of possibilities to future work. One of them would be to use scene change detection algorithms to better choose the key-frames showed in the timeline. This would facilitate users to recognize scene to be reviewed.

7. ACKNOWLEDGMENTS

FAPESP, CAPES, CNPq, FINEP, and MCT/CTIC.

8. REFERENCES

[1] F. Arman, R. Depommier, A. Hsu, and M.-Y. Chiu. Content-based browsing of video sequences. In *ACM Multimedia*, pages 97–103, 1994.

[2] M. G. Christel and N. Moraveji. Finding the right shots: assessing usability and performance of a digital video library interface. In *ACM Multimedia*, pages 732–739, 2004.

[3] C. Crockford and H. Agius. An empirical investigation into user navigation of digital video using the vcr-like control set. *Intl.J.Man-Machine Studies*, 64(4):340–355, 2006.

[4] S. M. Drucker, A. Glatzer, S. D. Mar, and C. Wong. Smartskip: consumer level browsing and skipping of digital video content. In *CHI*, pages 219–226, 2002.

[5] W. Hürst and K. Meier. Interfaces for timeline-based mobile video browsing. In A. El-Saddik, S. Vuong, C. Griwodz, A. D. Bimbo, K. S. Candan, and A. Jaimes, editors, *ACM Multimedia*, pages 469–478. ACM, 2008.

[6] H. ji Lee and J. hwan Kim. Apparatus and Method for Searching TV program. *Patent USPTO Application*, Patent(20080250451), 2008.

[7] J. Kim, H. Kim, and K. Park. Towards optimal navigation through video content on interactive tv. *Interacting with Computers*, 18(4):723–746, 2006.

[8] NCL.org. Ginga-NCL tools - Ginga-NCL Virtual STB - Linux Virtual Machine for VMWare, including Ginga-NCL C++ v. 0.10.1 rev.21. *http://www.ncl.org.br/ferramentas/fedora-fc7-ginga-i386.zip*, 2009.

[9] J. Nielsen and R. Molich. Heuristic evaluation of user interfaces. In *ACM CHI'90*, pages 249–256, 1990.

[10] R. F. Rodrigues, M. F. Moreno, and L. F. G. Soares. Presentation control of declarative applications in interactive digital tv system receivers. In *Brazilian Symp. Multimedia and the Web*, pages 184–192, 2006.

[11] L. F. G. Soares, R. F. Rodrigues, and M. F. Moreno. Ginga-NCL: The Declarative Environment of the Brazilian Digital TV System. *J. Brazilian Computer Society*, 12(4):37–46, 2007.

[12] E. Tulving. Episodic and semantic memory. In E. Tulving and W. Donaldson, editors, *Organization of memory*. Academic Press, 1972.

[13] M. Urbiola and Y. Martha. Capas neurológicas para la construcción del aprendizaje. *http://www.padhia.com.mx/noticias/revista_una.php?id=37*, 2008.

Experimental Characterisation of an Open Out-of-home MHP-based DVR Service

Philip Leroux
Ghent University-IBBT-INTEC
G. Crommenlaan 8 bus 201,
B-9050 Gent, Belgium
philip.leroux@intec.ugent.be

Kris Vanhecke
Ghent University-IBBT-INTEC
G. Crommenlaan 8 bus 201,
B-9050 Gent, Belgium
kris.vanhecke@intec.ugent.be

Klaas Roobroeck
Ghent University-IBBT-INTEC
G. Crommenlaan 8 bus 201,
B-9050 Gent, Belgium
klaas.roobroeck@intec.ugent.be

Chris Develder
Ghent University-IBBT-INTEC
G. Crommenlaan 8 bus 201,
B-9050 Gent, Belgium
chris.develder@intec.ugent.be

Filip De Turck
Ghent University-IBBT-INTEC
G. Crommenlaan 8 bus 201,
B-9050 Gent, Belgium
filip.deturck@intec.ugent.be

ABSTRACT

MHP (Multimedia Home Platform) is an open standard for interactive TV middleware that was designed by the Digital Video Broadcasting (DVB) alliance and is found on many commercially deployed set-top boxes (STBs). In this paper, a detailed system architecture is presented which enables customers of digital television providers to use their MHP-based DVR from an out-of-home location through their mobile phone. Building further on the openness of the MHP platform, special focus was on the overall scalability, security and extensibility of the system and the efficient and fast communication between mobile phone and set-top box. Technical challenges are imposed by the limited hardware resources that are available on most of the currently deployed STBs and by the delays that are introduced by the mobile communication networks. In order to characterise the system's performance, extensive throughput and scalability tests were performed on commercial DVRs. It is shown that the overall system performance meets the predefined requirements even under heavy load.

Categories and Subject Descriptors

H.5.1 [**Information Systems Applications**]: Multimedia Information Systems; D.2.11 [**Software Engineering**]: Software Architectures— *Domain-specific architectures, Service-oriented architecture (SOA)*

General Terms

Design

Keywords

Distributed Software design, Web Services, Java ME, MHP

1. INTRODUCTION

Due to the abundance of media, people have become more selective with regard to what they want to watch on television. This is confirmed by the growing sales of video recorders and the increasing popularity of services such as video-on-demand. In spite of their success, the current products and services on the market still provide no conclusive solution for a number of problems. The hardware capabilities of the devices are often limited (especially with respect to the CPU and RAM) and it is still necessary to be at home when programming the DVR. However, people often remember to record a program when outdoors and as such are forced to pay for view it on demand later on.

In this article, a generic architecture is presented that tackles this problem. It enables the user to configure his DVR remotely by using his mobile phone while taking into account the limited resources of both the set-top boxes and (most) mobile phones. This out-of-home recording may be provided as a service by the television provider, which implies that no additional configuration has to be done by the user nor does he has to purchase additional hardware.

The rest of the paper is organized as follows. In Section II, related work is discussed and we explain how our system architecture differs from other implementations. In Section III, we amplify on some requirements that our architecture at least should fulfill. Section IV provides an architecture overview, detailing each specific component and the interaction between these components. In section V, the interaction mechanism between the different components is described. In section VI results of the extensive performance tests that were performed on the implemented prototype are detailed. Finally, section VII states our conclusions.

2. RELATED WORK

Recently, several commercial operators started with offering similar out-of-home configuration options. Contrary to our system architecture, all these implementations are specifically targeting one specific architecture and one specific end user platform. Popular examples of these closed source implementations are DirecTV [1] in the USA and Sky [2] in the United Kingdom. Both implementations allow their clients

to configure their DVR through a web application or by using an iPhone application. Additionally, Sky customers can configure their DVR by sending an SMS. The work that is presented in this paper differs from these closed source implementations as it is an open and extensible framework that is built upon international standards and for more generic platforms, both for the mobile client side as on the STB. For the latter, all components were compliant with the MHP specification. MHP (Multimedia Home Platform) [3], is the collective name for a compatible set of middleware specifications developed by the Digital Video Broadcasting (DVB) alliance [4] and is currently deployed on many commercial STBs.

Other related work focuses on the out-of-home configuration of the DVR by deploying a home media center that is running at the customer premises [5]. Contrary to this work, we propose an architecture that will require no installation or configuration of additional components at the user's houses. Instead, a central deployable component with a general interface to the operator's management and payment components is introduced.

3. FUNCTIONAL REQUIREMENTS

After evaluation of related implementations and patents, it was defined that our architecture should enable the following functional requirements:

- By using his mobile device, the user should be able to securely login to the system and select a DVR from a list of DVRs that are linked to his account.
- An Electronic Program Guide (EPG) can be retrieved and browsed on the device.
- Based on the program selection in this EPG, the user can send a recording request to the DVR.
- Furthermore, a list of all scheduled records and recorded programs on a specific DVR can be retrieved.
- Scheduled records from this list can be deleted.
- Finally, the client application can also retrieve capacity information from every DVR linked to his account.

The main quality attribute requirements of the system, listed by decreasing importance, are: usability, performance, scalability, interoperability and modifiability.

4. ARCHITECTURE DESCRIPTION

One of the initial design decisions was the use of a central, management component that works as an intermediate layer between the user's mobile phone and STB. The use of such an intermediate layer was chosen, as direct communication between STB and mobile phone would introduce some inconveniences. First of all, in order to communicate with the STB, the STB's IP address needs to be known and it has to be accessible from outside the service provider's network where the STB is connected to. However, for security reasons, most service providers don't allow direct communication to their STBs from devices outside their private network. Second, the user would need to configure the mobile application to connect with that specific IP address.

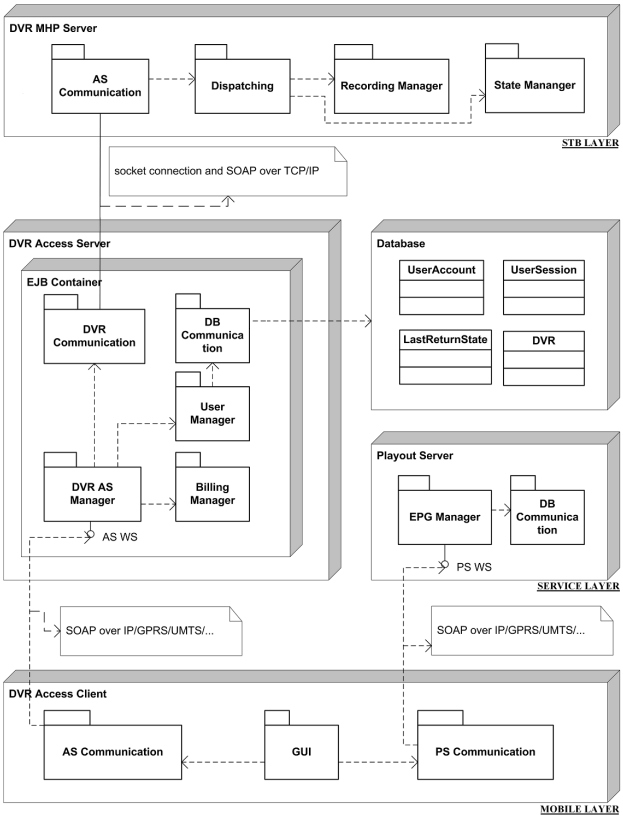
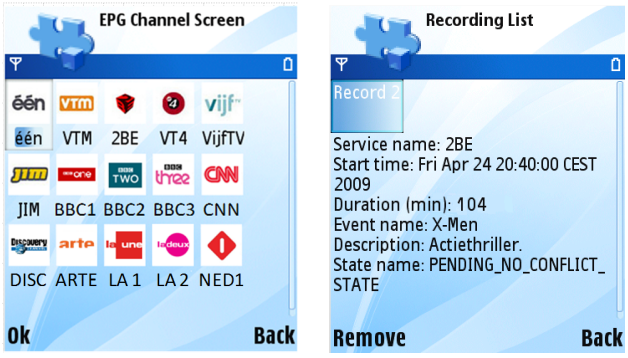


Figure 1: Low level system architecture.

As shown in the low level system architecture of Figure 1, three layers can be distinguished in the the general architecture: the client layer, the central service layer and the STB layer. The client layer consists of the client application which provides the presentation and communication to the user. The second layer contains the centralized processes. It consists of the DVR Access Server, which authenticates the user, manages all DVRs that are linked with this service and forwards all instructions to the correct DVR. The other components in the service layer are the Playout Server, which provides EPG data to the client application and a database component. This central layer is typically found in the backend of the service provider who manages all the STBs and most of the time also the digital video services. In the STB layer, an MHP application will run on the STB and process all incoming instructions. This is the DVR MHP Server application.

4.1 DVR Access Client

The main component (GUI) of the mobile application defines and initializes all graphical components and determines the flow of the user application. Two additional packages containing Web Service clients were added to communicate with the service layer. More specifically, *AS Communication* manages all communication with the DVR Access Server while the *PS Communication* packet is responsible for the Playout Server related communication. In both cases, the Simple Object Access Protocol (SOAP) was used. Once the user has logged into the system, he receives a list with all DVRs that are coupled with his user account. After selecting



(a) (b)

Figure 2: Screen shots of the client application ((a): EPG overview, (b): Recording information).

a specific DVR he can consult the program guide, requesting the current capacity of this DVR, consult and manage all scheduled recordings on this DVR or choose another DVR. Figure 2 shows two screen shots of the prototype client application. In order to target as many devices as possible, the mobile client was developed by using the Java Platform, Micro Edition (Java ME).

4.2 DVR Access Server

The DVR Access Server is the central component of the system and will be deployed in the back end layer at the server park of the service provider. This component is responsible for user authentication and distribution of the instructions to the various DVRs. The prototype that was tested during the performance analysis (detailed in Section VI), was developed by making use of Java EE (Java Enterprise Edition). It consists of several packages containing a number of stateless session beans, all providing different functionalities. One package is responsible for user authentication and management of sessions, while others retrieve database information or communicate with the DVRs. Two Web Services are designed, one to communicate with the DVR Access Client application and one with the DVR MHP Server.

4.3 Playout Server

The Playout Server is a Java EE application that is also part of the back end layer and whose Web Service can be called to retrieve all EPG data from a database. During the performance tests, live EPG data was used. Note that in operational systems, this component is already present in the backend of the service provider.

4.4 DVR MHP Server

Apart from the main module, that is controlled by the MHP Application Manager to start, pause and stop the application, four other modules are present. When the main application is started, the *AS Communication* module is initialized and starts listening to incoming messages from the DVR Access Server. When an instruction is received through this channel, it will be forwarded to the *Dispatching* unit for quick handling of the messages. All capacity related requests are then handled by the *State Manager* while all record related instructions are finally handled by the *Record Manager*. The prototype for the DVR MHP Server was developed using the MHP 1.1.2 Specification [3].

5. COMMUNICATION DETAILS

One of the main challenges for this use case, was the communication between the central service layer and the STB layer, due to the limited hardware resources of the STB. As a result, the communication between the DVR Access Server and DVR MHP Server applications is performed in an atypical manner. While the application client of the mobile layer uses the services of the service layer applications by addressing their Web Service, another approach has to be followed in this case. Due to the limited resources of the STB, no Web Service engine can run on the recording device.

In order to keep all advantages of a Web Service, following communication method was used: when an instruction arrives at the DVR Access Server application, the corresponding parameters are stored locally. An asynchronous socket message is sent to the DVR MHP Server, which triggers the device to call the server's Web Service and retrieve the necessary information to successfully execute the instruction. The result is then sent to the DVR MHP server using the same Web Service. To complete the process, the DVR MHP Server sends a socket message back to the DVR Access Server. The DVR Access Server will then evaluate the success of the operation and provide feedback to the user.

6. PERFORMANCE RESULTS

In order to evaluate the performance and scalability of the system, the developed prototype was subject to several tests. The prototype application was running on a commercial STBs that is currently used in many Belgian households. In cooperation with Belgian cable operator Telenet NV [8], this STB was running and fully functioning in their commercial cable network, which offers digital TV to hundreds of thousands subscribers. This had a major impact on the measurement results as on such commercial and fully functioning STBs, operators are typically running their own middleware processes in the background, already taken up a large portion of the STB's (limited) processing power. The main advantage of this approach, is that all results are representable for real life (out of the lab) situations. As all television channels were live data streams, automatic parsing of an up to date (external) program guide was performed in the Playout Server. During our tests, the Round Trip Time (RTT) of a request is measured in order to identify the numerous influences that cause a certain amount of delay.

In a first category of tests, only one request was sent to the DVR Access Server. The tests indicate that the data rate of the mobile communication channel, has an important impact on the RTT of a request. This is shown in Figure 3 where the average results of 10 tests are shown. For the UMTS connection, the network of the Belgian telecommunications operator Proximus was used. This UMTS network has a download link with an average data speed of 384 Kbps, which is a much lower data rate than a typical WiFi connection (802.11g) coupled with broadband internet access. As a result the RTT when using the UMTS network is significantly longer than when using the WiFi connection. Still, it is shown that even the heaviest operation *addRecord()* remains under the 5 seconds barrier. This 5 seconds bar-

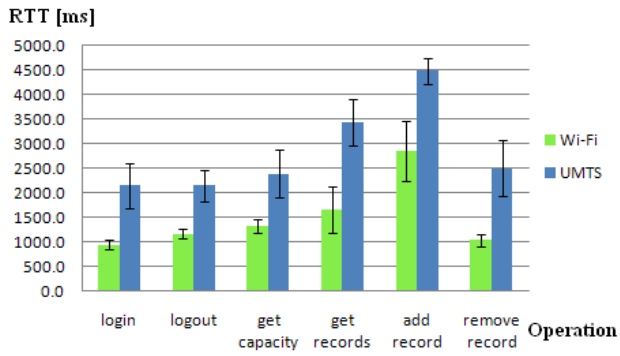


Figure 3: RTT (ms) for different mobile data transmission protocols.

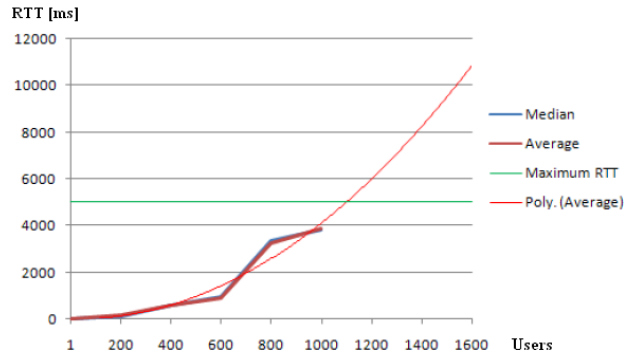


Figure 4: RTT (ms) of the login procedure as a function of the number of simultaneous login requests.

Table 1: Time difference of instructions (ms) depending on the recording activity.

Operation	Not Recording	Recording
get capacity	97,7	105,0
get records	87,1	138,4
add record	1429,0	1675,8
remove record	192,0	196,0

rier follows from a study performed by the Aberdeen Group where they wanted to investigate the correlation between the delay in a web application and the consumer's satisfaction of the service. Their report [7] states that the RTT should be limited to 5.1 seconds.

The impact of the recording activity is shown 1. These measurements were performed starting from the DVRCommunication module in order to filter out other influences. It is shown that the recording activity of the DVR will have a small negative impact on the execution time of all the instructions. The biggest absolute time difference (i.e. 246.8 ms) was measured when executing the *addRecords()* instruction. This is due to the fact that this relatively heavy instruction is imposed to the RecordManager of the MHP application, which is already responsible for the dispatching of the current recording.

Figure 4 shows a test using login requests. The RTTs were measured in every run, which differed in the amount of simultaneous login requests that were executed. A trend is added to predict the RTT using more requests. The green line represents the maximum RTT of 5 seconds considered acceptable by the user. The graph shows that the RTT is still acceptable when 1097 people log in at the same time.

7. CONCLUSIONS

In this paper, an open system architecture is described for the out-of-home configuration of a DVR by using a mobile phone. The system architecture requires no installation of additional components at the user's premises. Instead, a central component was installed at the Service provider's backend. This central component may interface with the Service provider's existing management system. One of the main challenges of the system was the implementation of efficient and reliable communication between different components with limited resources. This was implemented by combining synchronous and asynchronous communication calls. Additionally, decent user feedback is supported throughout the whole architecture. All functional requirements were fulfilled and the architecture was designed with usability, modifiability, interoperability, performance and scalability in mind. In order to evaluate the performance and scalability of the architecture, a prototype system was designed and thoroughly and successfully tested using live EPG data, a live transport stream and currently deployed commercial devices.

Acknowledgment

The authors would like to thank Bart Catrysse from Telenet N.V.[8] for his valuable input. Chris Develder is a Post-Doctoral Fellow of the Fund for Scientific Research, Flanders (FWO-V).

8. REFERENCES

[1] DIRECTV, `http://www.directv.com/`.
[2] Sky remote record service, `http://www.sky.com/`.
[3] Digital Video Broadcasting, *Multimedia Home Platform Spec. 1.1.2*, DVB Bluebook A068 Rev. 1, April 2005.
[4] Digital Video Broadcasting, `http://www.dvb.org/`.
[5] Shih-Hsiang Lo, Chi-Chia Lin, and Ming-Syan Chen, *Controlling Digital TV Set-Top Box with Mobile Devices via an IP Network*, IEEE Trans. on Consumer Electronics, vol. 152, number 3, 2006.
[6] Digital Video Broadcasting, *Digital Recording Extension to Globally Executable Multimedia Home Platform*, ETSI TS 102 817 V1.1.1, September 2007.
[7] Bojan Simic, Aberdeen Group, *Monitoring the End-User Experience for Web Applications: What Gets Measured, Gets Done*, January 2008.
[8] Telenet N.V., `http://telenet.be/`.

Monitoring of Hypertensive Patients through Conventional Medical Devices Integrated to the Brazilian Digital TV

Vandermi J. Silva and Vicente Ferreira de Lucena Júnior
Universidade Federal do Amazonas
Av. Rodrigo Otávio Jordão Ramos, 3000. Manaus-AM, Brazil
55-92-3305-4671
{vandermi, vicente}@ufam.edu.br

Jomara Bindá
Universidade Estadual do Amazonas
Av. Djalma Batista, 3578 Manaus-AM, Brazil
55-92-3214-5774.
jomarabinda@gmail.com

ABSTRACT
This paper presents an implementation of a healthcare monitoring system integrated to the model of Digital TV adapted in Brazil (SBTVD), which provides the convergence among wireless sensor network, mobile devices, the World Wide Web (WEB) and embedded systems, through the connection of these devices to a central computer called healthcare gateway. This gateway manages the service requests and access via Digital TV (DTV) and mobile devices like cellular phones. Conventional sensors capture the patient's body temperature, pressure, heartbeat and room temperature which are sent to the gateway. Then a reactive software agent based on the behavior, verifies the status of the sensors and dispatch a message to a DTV and/or a cellular phone previously registered in the system via WEB interface. The impact of this new concept will be very high on the Brazilian population due to the great number of people suffering from hypertension and the high level usage of digital TV and mobile devices can help health teams in monitoring the patient.

Categories and Subject Descriptors
C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design – network communication, wireless communications.

General Terms
Design, Experimentation, Management

Keywords
Healthcare Systems, Medical Devices, Brazilian Digital TV, Wireless Sensor Network

1. INTRODUCTION
According to IBGE [7], over 90% of Brazilians living in urban areas, have access to television, and certainly that percentage should be reflected in the new DTV system, when it is implemented in Brazil in its entirety. Coupled with this fact, in [7], the IBGE also shows in its research that the number of elderly people in Brazil accounts for more than 19 million people aged 60 years and over. This is 11% of the population, which continues to grow and should reach 30 million over the next 20 years, or almost 13% of the population at the end of this period.

This leads us to believe that demands for specialized health services will grow in proportion to the growth of the elderly population, affecting healthcare in hospitals and nursing homes, while with the increasing age of the citizens, their chances of contracting diseases will also increase.

On the other hand, the market needs applications that automate the processes for monitoring patients remotely to allow quality care and reduce queues at the clinics. So this work can help to improve the quality of life of patients and connect devices with conventional technology and the new digital TV recently implemented in Brazil.

2. SCENARIO OF THE APPLICATION
To better illustrate how the work was done, a scenario simulating a situation in which a patient is monitored through the system, will be presented.

Mary and Joseph, an elderly couple, aged 75, live alone, have heart problems and always forget to take the medication prescribed by their doctor. Once a month, Mary and Joseph are examined at the hospital where they receive medications that should be taken daily to facilitate their lives. Their son setup an automated control system integrated with the DTV and the GSM network to monitor their health status.

The system collects data from a device equipped with a pulse oximeter, a temperature sensor and a blood pressure tester that sends information to a central computer at home equipped with a message alert system.

We can see in Figure 1 (B) this scenario based on an automated home, which has a local network integrated with an Internet router and a gateway to manage the devices in the house. Thus,

taking advantage of this infrastructure, we insert the sensors and medical devices in the network using wireless networks connecting the gateway.

The equipment is connected via Local Area Networking (LAN) and Wireless Local Area Networking (WLAN) and is used to transmit SMS messages via cellular phones integrated at gateway. Completing the scenario, Figure 1 (A) represents the mobile system receiving the warning messages via cellular network and the registration system accessing the data from home via WEB.

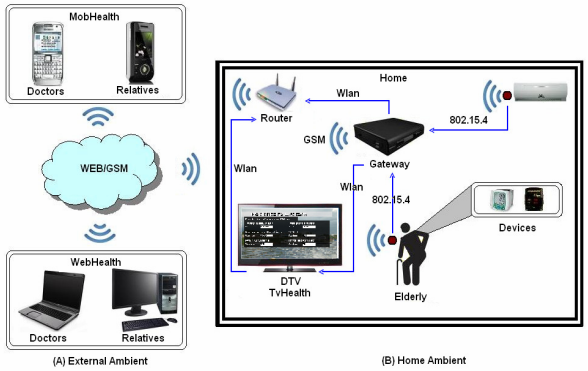


Figure 1: Scenario for a Monitored Residence

3. RELATED WORKS

Several applications can be implemented from the scenario presented, such as those of tracking patients using Radio Frequency Identification (RFID) access control and location of people, as shown in [3]. Using RFID the authors present a nursing home monitored via the central computer, which processes the data and sends it via wireless to the doctor's handheld and caregivers for the elderly. In this case, the authors used RFID tags to identify patients in different areas of their homes such as doors, corridors and rooms.

Another work presented in [11] provides a ubiquitous environment healthcare integrating a Zig Bee sensor network through the OSGI framework Digital TV device using the MHP standard. In this work the authors present the architecture and define a specific protocol to convert Zig Bee and virtual UPnP devices by mapping IP addresses to ports of the devices. Thus devices such as Zig Bee glucose monitors are monitored within the home environment and outside environment through the WEB.

In [8], a framework for intelligent services healthcare is presented, where sensor fault detectors send information to a system decision based on Ontology and the Semantic WEB. Depending on the analysis data, the system triggers an alarm in the house such as a siren sending the information via the WEB to the healthcare team. In this work the authors validated the architecture of the system through a scenario of a falling detection device.

In [4] the authors show a framework for a system of healthcare wireless using smart phones as the hub of sensors connected to the patient's body through a Body Sensor Network (BSN). The system detects the patient's vital signs and then sends the signals to a server that processes the data and sends it to the health center. The objective of this work is to ensure safety and Quality of Service (QoS) transactions in the collection and transmission of data using the BSN.

In our work we focus on the development of software architecture in X86 platform that encapsulates the virtual machine JAVA and JADE, installed in a Linux distribution, an application to monitor blood pressure, heartbeat and the patient's temperature. Moreover, our gateway has a WEB server, an application for digital TV and a cellular phone connected to provide SMS for doctors, nurses and relatives of patients being monitored

We note that in most studies found, we find basically the use of a gateway, coupled to a sensor network, although only one presented in [8] of them had a layer of decision to help the system, using anthologies and semantics. The main difference among these works and our works is a new approach applied in our gateway to involve the multi agents ambient and Digital TV for the monitoring of patients' hypersensitivity. Thus we can monitor the events in the residence of the patient and help the health professionals take decision.

4. PROPOSED SYSTEM

The architecture developed in this study used the specification of the Brazilian DTV middleware presented in [2], specifically the layer Ginga-J, which allows of the Xlet Java applications will be executed.

Our proposed system is divided into hardware devices and software represented respectively by a central computer gateway, conventional medical devices, mobile phone and DTV.

4.1 Hardware

The residential gateway hardware uses the X86 architecture and is composed of an Intel Atom Dual-core processor, a motherboard (AsRock A945GC model) equipped with Intel chip-set network card gigabit Ethernet and sound cards and integrated video, 1 GB (gigabyte) of RAM and a device storage (pen drive), 8 GB, used to install the Debian Linux operating system, applications and file system.

The processing unit has a interfaces for connection sensors and a serial bus for communication with the central computer gateway to transmit the data. It also has a unit that provides the battery voltage (VBAT) that is divided between sensors and processing unit to allow the unit to be self-sufficient and portable.

The interconnection among the devices for data collection and the interface of the gateway was made through a circuit board developed in the laboratory of Electrical Engineering at the Universidade Federal of Amazonas as shown in Figure 2.

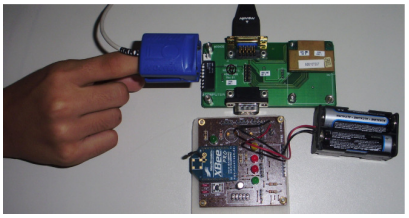


Figure 2: System for Data Collected with Pulse Oximeter

In this work we used a pulse oximeter developer kit from NONIM Medical [10] modified to include a Zig bee module shown in Figure 2.

The photo detector receives the light from the sensors and detects the difference between the light that was absorbed by hemoglobin molecules. The heart rate is calculated from measuring the time interval between the detection of the peaks of the waveform of the infrared.

4.2 Software

The system architecture has five layers, namely application, decision, JAVA Virtual Machine (JVM), Operating System (OS) and Communication (COMM). The following will show in details each layer of the system and their functions. The application layer is responsible for integrating all applications developed, such as WEB applications, mobile and DTV. This layer uses the JVM and WEB Server installed in the operating system, though the architecture which allows applications to be developed without the use of the JVM, for example applications developed in C running directly in the operational system.

The decision layer is implemented using the JVM through Java Agent Development Framework (JADE) presented in [1] that is responsible for managing applications messaging and status updates. This is done through the use of behavior in pre-established agents which are capable of deciding whether to send SMS with the changes found in sensors.

The decision layer is instructed to monitor the heartbeat received from the pulse oximeter, check the temperature received from the sensor and check the temperature and pressure measured by the patient through the blood pressure tester. If the parameters are normal, agents continue monitoring and do not send messages. However if the parameters are abnormally exceeded, the agent sends a message to a second agent that triggers an SMS to mobile phone users and sends an alert to the Set Top Box which is presented on the DTV screen.

The protocol messages of agents have the format of the Agent Communication Language (ACL) defined in [5] and have a number of predefined fields, among them cited here: The sender of the message, the list of receivers, the communicative intention, the content and the ontology.

In the SO layer, there is the Linux operating system, customized for the operation of the gateway to healthcare, which has been configured in the JVM, WEB Apache Server, PHP, JADE, Zig Bee, and RXTX API for USB support, RS232 and SMSLIB API support for sending SMS. Figure 3 presents the division of the modules in high level, through all the layers presented.

In the application layer, the modules *DataRegister*, *HcMonitor*, *SendSMS* and *DTVAlarm* were implemented by fully sharing the same Extensible Markup Language (XML) file, located in the root directory of the gateway. This XML file is monitored by *HcMonitor* and classes *Oximeter*, *Pressure* and *Temperature*, responsible for collecting data from sensors.

Through the registration of users and sensors, the module generates the XML *DataRegister* collects initial information from patients and *HcMonitor* module does the monitoring and updating of such information. Figure 4 shows the architecture.

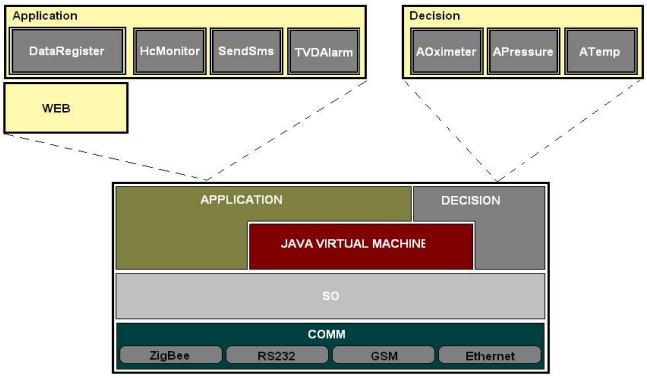


Figure 3: The General Architecture Dived by Layers

5. Implementation and Results

Initially the devices chosen were oximeter, pressure tester and temperature sensors. Then the application scenario was specified and the identification of modules of the system was made. At this stage, the UML [6] was used, initially with generated use cases, subdivided into cases of specific uses for each module.

The next step was to separate the decision layers and applications. Then we developed applications that were separate and independent. An example of this class can be seen in Figure 4. In this diagram, the data package has two classes, the Oximeter class and the Pressure class; both classes access the serial ports through a listener that runs on thread controlled by the Java Virtual Machine

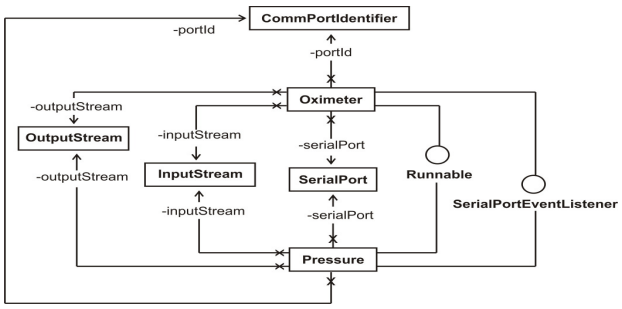


Figure 4: A Fragment of the Oximeter and Pressure Class

The module gateway was installed on the JADE framework, to implement the agents, the Apache WEB server and the PHP interpreter scripts to develop the interface configuration for the data gateway. Then a folder structure to store files in XML format was created. The data of patients and the sensors connected to the gateway offer access to these files and other modules of the system.

The file structure consists of one root tag that has <*data*> with sub-nodes <*personal_data*> and <*sensors*> for each device. They are all the attributes and status attributes for the collection of data from their respective sensors.

The WEB module consists of a WEB server installed in the gateway with access to database in XML. This database stores the data from the sensors and personal data of the patient, doctor and relatives.

The DTV module was developed using the Java TV API and consists of a module that runs on DTV and communicates with the gateway over the local network in two ways, through the remote control pressed by the user or by receiving the warning signal provided by the monitoring system.

The application developed for DTV, has three packages called *comm, main* and *ui*. The package *comm* is responsible for communicating with the class through a sensor interface, which collects data from the XML file and makes them available to the *ui* package through the classes *OximeterBox*, *PatientBox, PressureBox* and *TemperatureBox*. The result can be seen in Figure 5.

Figure 5: The TV Healthcare Application

The mobile module is an application developed in JAVA Micro Edition, embedded in the phone. This module is responsible for opening a connection via the GSM network and verifies the data written to the gateway. In this way you can track the current status of the sensors using the phone.

The verification of the status of the sensor made by phone occurs through a reading of the XML file stored in the gateway which is constantly being updated. Figure 6 shows an application example.

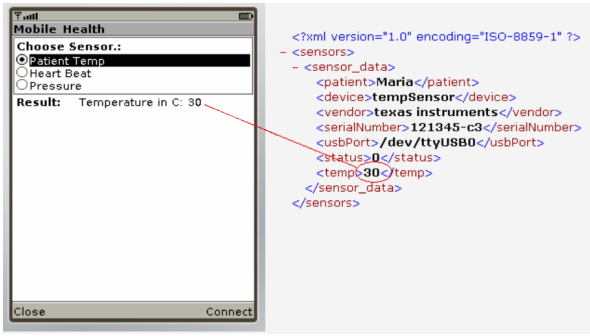


Figure 6: Mobile Phone Interface

6. Conclusions and Future Works

The results of this study were the applications themselves developed to validate the research and integration of existing technologies with the Brazilian DTV. Studies in this area in Brazil are in their early stages and efforts to improve patient care through the use of technology are increasingly growing.

During the implementation of the work, difficulties have been overcome, mainly in the integration of devices from different manufacturers. The prototype was tested on a gateway compatible with the standard Brazilian DTV and for future work we intend to test the system on a Set Top Box, using the mobile telephone network as an application return channel. The experiments will be done with other devices and other technologies using the proposed architecture, such as devices for measuring glucose, balance among others. Thus, we hope that this work will contribute to the health teams in Brazil at the requirements for mobility and quick service.

7. Acknowledgments

We would like thank the CAPES, CNPq, FAPEAM, CETELI, and the DCC-UFAM that provided the support for the implementation of this paper.

8. References

[1] Baima Remis Lima. Realization and Optimization of an Agent-Based Test Resources Manager. Master Thesis of the Universität Stuttgart , 2009.

[2] L. F. G.Soares, R. F . Rodrigues and M. F. Moreno. Ginga-NCL: The Declarative Environment of the Brazilian Digital TV System. Journal of the Brazilian Computer Society, 13(1):37-46, March 2007.

[3] J.M. Corchado, J Bajo, and A. Abraham. Gerami: Improving Healthcare Delivery IN Geriatric Residences. IEEE Intelligent Systems, 23(2):19-25, 2008. http://doi.ieeecomputersociety.org/10.1109/MIS.2008.27.

[4] S. Dagtas, G. Pekhteryev, Z. Sahinoglu, H. Çam, and N. Challa. Real – time and secure wireless health monitoring. Int. J. Telemedicine Appl., 2008:1-10, 2008. http://dx.doi.org/10.1155/2008/135808.

[5] FIPA. Foundation for Intelligent Physical Agents. http://www.fipa.org/repository/standardspecs.html. Access in 2009 October.

[6] Fowler Martin and Scott Kendall, UML Distilled. A Brief Guide to the Standard Object Modeling Language. Addison-Wesley, 2000.

[7] IBGE. Pesquisa Nacional por Amostra de Domicilios. http://www.ibge.gov.br, 2009 October. Access in 2009 November.

[8] M. Valero, L. Vadillo, A. Penalver, I. Pau, E. Gago, M. Martin, and M. Gonzalez y Eloy Portillo. An implementation framework for smart home telecare services. In Future Generation Communication and Networking (FGCN 2007), volume 2, pages 60–65, Dec. 2007.

[9] Morris Steven, Anthony Smith-Chaigneau. Interactive TV Standards A guide to MHP, OCAP and Java TV. Elsevier, USA 2005.

[10] NONIN Medical. Oem-iii-module pulse oximeter. http://www.nonin.com/OEMSolutions/OEM-III-Module. Access in 2009 October.

[11] W.-W. Lin and Y.-H. Sheng. Using osgi upnp and zigbee to provide a wireless ubiquitous home healthcare environment. In UBICOMM '08: Proceedings of the 2008 The Second International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, pages 268–273, Washington, DC, USA, 2008. IEEE Computer Society.

Multi-user Interactive TV: the next step in personalization

Ray van Brandenburg
TNO ICT
Brassersplein 2
2621 CT Delft, the Netherlands
ray.vanbrandenburg@tno.nl

M. Oskar van Deventer
TNO ICT
Brassersplein 2
2621 CT Delft, the Netherlands
oskar.vandeventer@tno.nl

Georgios Karagiannis
University of Twente
Drienerlolaan 5
7522 NB Enschede, the Netherlands
karagian@cs.utwente.nl

Mike Schenk
TNO ICT
Brassersplein 2
2621 CT Delft, the Netherlands
mike.schenk@tno.nl

ABSTRACT

In the past few years there has been an increasing trend towards personalization in the TV world. IMS-based IPTV is a good example of a highly personalized IPTV architecture, featuring an advanced identity management subsystem. This article studies a next step in the personalization of the television experience: concurrent use of TV services that are supported by the IMS-based IPTV system. That is, multiple users using the same television set at the same or at different times, where each user has personalized interaction with the service and content.

Our analysis of IMS-based IPTV use cases shows that current architectures were not designed with concurrent use in mind. However, we demonstrate that the combination of concurrent use and personalized TV services can yield interesting and viable use cases in the areas of interactive game shows, personalized electronic program guides and channel lists, and other.

Finally, an analysis of the IMS-based IPTV system and architecture shows that it has all the ingredients to implement these new concurrent TV use cases, and that the main challenges will be in the area of usability. The article concludes that personalization and concurrency are not contradictory for television services, neither from a use case perspective, nor technologically. In particular, the IMS-based IPTV system is able to facilitate an enhanced and personalized experience to concurrent TV users.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design – *network communications*

General Terms

Human Factors, Standardization, Design

Keywords

IMS, IPTV, personalization, multi-user, concurrent TV services, use cases, challenges.

1. INTRODUCTION

Concurrent use of a TV by several users has always been an important part of watching TV. A well known scenario is that of multiple family members or friends watching TV together at the end of the day. Another example is that of big sports events such as the Super Bowl in the US and the World Cup in Europe, which have become major social events where large groups of people get together and watch collectively.

These examples of concurrent use seem contradictory with the recent trend in the TV world towards personalization and with it the focus on a single user per TV; the collective experience versus the individual experience. The main goal of this paper is to analyze and discuss whether the IMS-based IPTV system and architecture can be used to facilitate a personalized experience to concurrent TV users. Moreover, this paper aims to verify whether the hypothesis that personalization and concurrency are contradictory, is valid. We will try to work out this issue both from a use case perspective and from a technological perspective.

This paper is organized as follows. Section 2 provides some background on the evolution of TV as a personal medium. It also briefly introduces (IMS-based) IPTV as enabling technology for more advanced forms of personalization. Section 3 analyses the use cases that form the basis of IMS-based IPTV to find out to what extent concurrent use has been taken into account. Section 4 proposes a new set of use cases that involve multiple users, personalization and concurrent use of TV services. It demonstrates that there are interesting and viable use cases in this area. Section 5 studies whether and how the IMS-based IPTV architecture can be used for the above-mentioned type of use cases. Section 6 concludes that personalization and concurrent use are not contradictory, neither for the use cases nor for the technical implementations.

2. THE EVOLUTION OF TV AS A PERSONALISED MEDIUM

This section provides some background on the evolution of TV as a personal medium. When TV was introduced more than fifty years ago, it was delivered over the air using antennas. In these early days of TV, identification was not possible; there was no way for TV broadcasters to know how many TVs or which users would be receiving their signal. The result of this was that TV was a free medium; broadcasters simply could not charge TV owners.

With the introduction of cable TV, a simple form of identification was introduced; the physical connection from TV to provider. It was now possible to offer larger bundles of television channels to households and (more importantly) charge them for access.

The introduction of scrambling of TV signals, such as the Conditional Access (CA) method found in DVB [1], provided TV broadcasters and content providers with a system to more effectively control subscriptions to television channels. By scrambling TV channels with unique keys and then distributing these keys to subscribers in different packages it was possible for providers to offer different service levels and differentiate between households in the same neighborhood. This marked the first time that TV became ‘personal’ in the sense that people could control which channels they received.

In parallel to these developments in the digital TV world, early forms of interactive TV were introduced. By using a separate return channel, not directly linked to the TV access technology

used, customers were for the first time able to interact with TV shows. In the beginning, this return channel was the telephone line, later SMS and the internet became the return path of choice. Present day examples include SMS voting on popular TV shows such as Idols and Big Brother.

While subscription TV and interactive TV have been separate developments towards personalization for quite some time, the two trends are growing towards each other in IPTV. The distinguishing feature of IPTV is that both the subscribed content delivery and the interactivity are provided over the same IP network. This integration enables more advanced use cases using user identification instead of terminal identification. Social TV [2] applications in particular make use of these new features. These applications integrate the television experience with social interaction, such as text- and video chat, content recommendations between users and sharing status information between users sitting at different TVs. A good way to describe Social TV is "Watching apart, together".

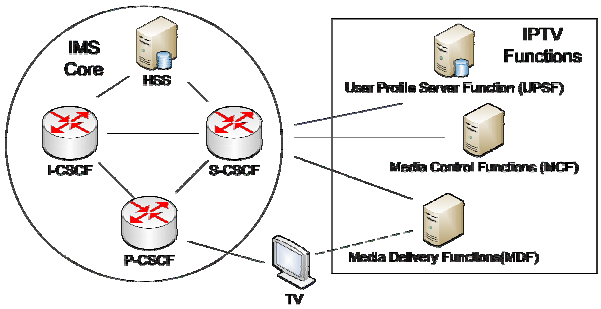


Figure 1. IMS as an intermediary between the TV and the IPTV functions

The rest of this paper focuses on a particular IPTV architecture, called IMS-based IPTV [3]. The reason for this is that it features both an extensive identity framework that supports identification of individual users and a session-control framework for advanced forms of interactivity. For an overview of the structure of IMS-based IPTV, see Figure 1. At the center of the architecture is the IMS core, the architectural framework that provides the session control functions and the authentication procedures. The IMS core acts as an intermediary between the IPTV functions and the TV. All these aspects make IMS-based IPTV intrinsically suited for personalization of the TV experience.

3. CURRENT FOCUS ON SINGLE USERS

IMS-based IPTV is standardized by the TISPAN body of the European Telecommunications Standards Institute (ETSI). The blueprint for the architecture lies in document TS 181 016 [4] (which it shares with the Integrated IPTV architecture). By analyzing this document, we can get an idea of the focus of the architecture. Which types of use are included and which are considered most important? The central question of course is, to what extent has concurrent use been taken into account?

IPTV use cases can be divided into three different categories: classic TV use cases for arbitrary numbers of users (not dependent on the number of users), use cases for multiple users at different TVs (one per TV) and the concurrent use category of multiple users at the same TV (either simultaneously or at different times).

As one looks at the use cases offered in TS 181 016, it immediately becomes obvious that the majority of the use cases fall in the first category of arbitrary number of users. Examples are traditional broadcast TV and the electronic program guide (EPG), but also use cases such as Video-on-Demand, Personal Video Recording and Pay-Per-View.

Use cases in the second category of multiple users at different TVs are mostly use cases that fall under the 'Social TV' banner such as presence, Watching Apart Together, recommendations from other users and sharing the remote control.

The third category, use cases for several users of the same TV, is empty; no real use cases or services were found that are explicitly meant for concurrent use. It should be noted that if one looks very good, a single line can be found that does refer to the multi-user situation. This line, which is related to the presence use case, is the following: "The interactive IPTV solution shall allow multiple users in front of one TV-SET to communicate their status". What this means, is that out of tens of use cases, and hundreds of requirements, there is only a single sentence devoted to concurrent use. Even more remarkable is the fact that in the final stage 2 document (TS 181 027 [3]), that describes the IMS-based IPTV architecture itself, there is no mention of this requirement and the feature is not explicitly implemented.

Based on this analysis, it is safe to say that the IMS-based IPTV solution, while enabling many personalized services for individual users, has not been designed with multiple users and concurrent use in mind.

4. CONCURRENT-USE USE CASES

The observation that the above-mentioned use cases for a personalized TV experience focus on single individual users, raises the question whether personalization and concurrent use are contradictory. To test this hypothesis, we have tried to come up with the opposite: use cases that combine both aspects; personalized use cases for multiple users sharing the same TV. In a number of brainstorm sessions we have found several of such use cases, among others the following three:

1) Interactive TV show: TV game shows have been around for many years. One of the possible reasons game shows have been such a success is the ability for viewers at home to "play along" by thinking about the questions and discussing them with friends and relatives. Although some interactive TV products have responded to this with applications in which users can play along using their remote control, these products are limited to one user per TV. Having multiple viewers sitting on the couch next to each other, competing against each other, can add both a social element and a form of competition to it, making it a richer experience. An example which proves people like to play these types of games together is the success of the PlayStation game Buzz! , which is basically an offline TV quiz. Total sales for Buzz! games have reached 6 million as of May 2008 [5]. This use case is good example of how multiple users enjoying a TV application together can enrich the experience, making it more interesting than a single-user experience.

2) Personal EPG/Channels: In recent years, both in the US and Europe, the number of TV channels has skyrocketed. This makes it increasingly harder for customers to make a selection; users are overwhelmed by choice. To add to this problem, different members of the same household often have different tastes, each

family member subscribing to different channels. This results in a TV that contains hundreds of channels of which every member only watches a very small subset [6][7]. A useful multi-user application would be a system in which every family member is able to choose and order his own list of channels. When this user turns on the TV, he only sees his favorite channels, in the order of his choosing. When two or more members are simultaneously logged in, the subscribed channels are added to each other. An additional possibility is for adult-only channels to automatically disappear when a child logs in.

The personal EPG/Channels use case is an example of a use case meant for users sharing a TV primarily in a time-share manner.

3) Multi-user Social TV: Social TV applications allow users at different TVs to chat and enjoy TV shows together. Taking multiple simultaneous users into account brings additional opportunities and requirements to the concept of Social TV. One of those requirements is privacy; it is important that users on the other side of the TV can see that you're not alone, or they might not send those very private chat messages (and ensuing embarrassing situations). This shows that taking multi-user scenarios into account is not only a nice extra but also a necessity to ensure privacy.

These three use cases (and others) show that there are interesting use cases that involve both personalization and concurrent use of one TV, hereby proving that the hypothesis of concurrent use and personalization being contradictory is not valid. Even more, they show that in some cases concurrent use and personalization are not only non-contradictory but require each other; concurrent use increasing the need for personalization (multi-user EPG) and personalization requiring multi-user scenarios being taken into account (multi-user social TV).

5. IMS-BASED IPTV IS SUITED FOR CONCURRENT USE

In the previous section it was demonstrated that, from a use case perspective, personalization and concurrent use are not contradictory. Even more so, they can enrich each other. This section will deal with the same question but from a technological perspective. As has been shown in section 3, IMS-based IPTV has not been developed with concurrency in mind, which poses the question if it is able to support multiple simultaneous users at all. More specifically: "does IMS's identification mechanism support multiple simultaneous users at the same terminal?" In order to answer this question, we first need to get an idea of the kind of requirements the presented use cases introduce.

The one thing all three use cases share is the need for identification; a way to uniquely identify the different users sitting at the TV together.

Roaming would make each of the three use cases even more useful; roaming in the sense that users should not be bound to specific terminals but should be able to log in at any TV within the network. This could mean TVs in their bedroom or basement but also TVs at a neighbor's or a friend's house. (It should be noted that roaming in this context is different from the same term when used in the mobile telephony world, where it refers to the process of switching from one carrier to another). The concept of roaming also introduces a third requirement, the fact that two users of different subscriptions should be able to watch TV together on the same terminal. All of these requirements have been turned into four different user/group configurations, visualized in Figure 2.

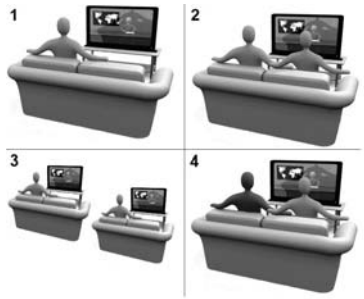


Figure 2. Four different user/group configurations

5.1 Identifiers in IMS

The IP Multimedia Subsystem (IMS) [9], on which IMS-based IPTV is based, has its origins in the mobile telephony world. This becomes evident when one looks at its identity framework, which is heavily based on the authentication applications stored on Universal Integrated Circuit Cards (more commonly known as SIM cards). There are both advantages and disadvantages to IMS's origin in the mobile telephony world. On the one hand, a mobile telephone is the perfect example of personalization. On the other hand, the mobile telephony world does not feature an analogue for the concept of multiple users simultaneously sharing a terminal. Therefore the various identification mechanisms (which are designed for this single-user world) might not be suited for the concurrent use scenario. This is what will be checked during this section.

At the core of the IMS identification mechanism are two distinct identifiers: the IMS Private Identifier (IMPI) and the IMS Public Identifier (IMPU). The IMPI can be compared to the IMSI found in GSM/UMTS. It is a private identifier in the form of a SIP (Session Initiation Protocol) URI that is bound to a specific subscription. The IMPI is used for authentication purposes and is only known to the UICC and the operator's authentication centre.

The IMPU is a public identifier in the form of a SIP URI or TEL URI that can be compared to the telephone number (MSISDN number) in the telecom world. It is possible for a user to have multiple IMPUs associated with a single IMPI (just as a user can have multiple phone numbers). It is also possible for one UICC to have multiple IMPIs (with their associated IMPUs). It is important to note that the IMPU plays no role in the authentication process, which is the domain of the IMPI.

Figure 3 demonstrates that through the use of IMPIs and IMPUs, IMS's identification system supports a wide variety of configurations in which users can have both individual as well as shared identifiers. The question now is: is it possible to simultaneously register two IMPIs (possibly of different subscriptions) on the same terminal? Fortunately, the IMS standards are very clear that this is indeed supported. What this means, is that in theory, the concurrent use scenario is feasible.

However, before we can conclude by saying IMS-based IPTV supports concurrent use, we first need to look at the more practical aspects of IMS's identification system. The main issue here is: how to get two IMPIs from different households on the

same terminal? In order to answer this question we need to look at IMS's authentication system. This system relies primarily on a UICC based mechanism called IMS AKA (IMS Authentication and Key Agreement). In order to provide authentication for terminals that do not feature UICC slots, two other mechanisms, SIP Digest and NBA, are also provided.

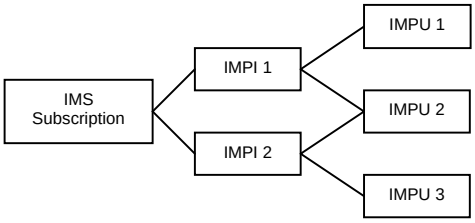


Figure 3. IMPI/IMPU allow for different configurations

5.2 Authentication in IMS

IMS's main authentication method, IMS AKA, can be compared to similar AKA mechanisms in GSM and UMTS. It is a challenge response mechanism that uses a combination of IMPI and a shared secret K that is stored in the UICC and in the operator's authentication centre.

As already discussed, IMS allows for multiple IMPI to be placed on a single UICC. If we now assume that each user is identified by a single IMPI, there are two options for a family. The first is to give each family member a separate UICC containing their personal IMPI. The second option is to have a single family-UICC containing the IMPIs of all family members.

Now consider the four user/group situations described earlier and shown in fig. 3. If the family members each have their own UICC, there is a problem when two members want to watch TV together: the TV/STB needs two slots. The problem gets worse when there are three, four or more family members watching TV together.

The second option, with a family-UICC, is also not ideal. In this case there are problems when different family members want to watch TV at different locations (either in their own house or at other households). Since there is only one UICC, only one TV can be used simultaneously.

The main problem with IMS AKA is that the use of physical UICCs does not scale. Fortunately there are two more authentication methods. The first of those, NASS-bundled Authentication (NBA) is not really an option. NBA works by linking an IMPI to a physical line identifier. This means that an IMPI can only be used from a specific access line. Needless to say, this is very inflexible and makes roaming impossible.

The final authentication mechanism is SIP Digest. SIP Digest is based on HTTP Digest, which is used for website authentication. Instead of being based on a UICC, SIP Digest is built around a combination of IMPI and password. When a user registers in the IMS core network, he provides his IMPI. His password is then checked using a challenge-response mechanism.

Note that SIP Digest is intrinsically less secure than IMS AKA [8], so an additional layer of security (DNS-SEC, https, SSL) may be needed for its use. For most internet applications, Digest-based security suffices, and we believe it is also sufficient for access to IPTV services. Also, SIP Digest requires a user-friendly system to enter username and password, e.g. using a cookie or an RFID to store this information

What we have shown in this section is that, although there are some usability problems stemming from the way IMS's authentication system works, in theory, there are no problems with IMS-based IPTV's identity framework supporting multiple concurrent users. This means that just as there is no contradiction from a use-case perspective between personalization and concurrent use, there is also none from the technological side.

6. CHALLENGES ON USABILITY

This article shows that up until now concurrent use has not been explicitly taken into account in the design of new TV architectures such as IMS-based IPTV. In this paper we show that personalization, which is playing an increasingly prominent role in the TV world, is not contradictory with concurrent use and by introducing a number of interesting multi-user use cases we demonstrate that, in fact, the two concepts can complement each other. Finally, we have shown that, even tough it was not designed with concurrent use in mind; with a few minor issues IMS-based IPTV does allow such use cases to be implemented. Therefore it can be concluded that the IMS-based IPTV system can facilitate an enhanced and personalized experience to concurrent TV users.

Future work will focus on overcoming these issues and developing and evaluating a prototype that will implement the concept of personalized concurrent use.

7. REFERENCES

[1] Digital Video Broadcasting, "Support for use of scrambling and Conditional Access", ETR 289 v1

[2] E. Boertjes, J. Klok, O. Niamut, M. Staal, "ConnecTV: Share the experience", in *Social Interactive Television*, P. Cesar, D. Geerts, K. Chorianopoulos: IGI Global, 2009, pp. 187-221

[3] ETSI TISPAN, "IPTV Architecture: IPTV functions supported by the IMS subsystem", TS 182 027 v2.0.0

[4] ETSI TISPAN, "Service Layer Requirements to Integrate NGN Services and IPTV", TS 181 016 v3.2.0

[5] Eurogamer, "PS3 has outsold Xbox 360 in Europe", eurogamer.net, Oct. 7, 2008. [Online]. Available: http://www.eurogamer.net/article.php?article_id=137142. [Accessed: Dec. 22 2009]

[6] Nielsen Media, "Average U.S. Home Now Receives a Record 118.6 TC Channels", nielsenmedia.com, Jun. 6 2008. [Online]. Available: http://www.mediabuyerplanner.com/entry/34086/us-homes-receive-a-record-1186-tv-channels-on-average/. [Accessed: Dec. 2 2009]

[7] RNW, "Over 6,500 TV channels in Europe, UK in top spot", rnw.nl, Oct. 17 2008. [Online]. Available: http://blogs.rnw.nl/medianetwork/over-6500-tv-channels-in-europe-uk-in-top-spot-poll. [Accessed: Dec. 3 2009]

[8] ETSI TISPAN, "3G Access Security for IP-based Services", TS 133 203 v8.5.0

[9] ETSI TISPAN, "IP Multimedia Subsystem: Stage 2", TS 123 228 v8.5.0

Design of an Architecture for Management of Context-Aware Narrowcasting

Klaas Roobroeck
Ghent University - IBBT
Gaston Crommenlaan 8 bus 201
9050 Ghent, Belgium
klaas.roobroeck@intec.ugent.be

Philip Leroux
Ghent University - IBBT
Gaston Crommenlaan 8 bus 201
9050 Ghent, Belgium
philip.leroux@intec.ugent.be

Tim Wauters
Ghent University - IBBT
Gaston Crommenlaan 8 bus 201
9050 Ghent, Belgium
tim.wauters@intec.ugent.be

Filip De Turck
Ghent University - IBBT
Gaston Crommenlaan 8 bus 201
9050 Ghent, Belgium
filip.deturck@intec.ugent.be

ABSTRACT

Each day we are dazzled by advertisements; surfing the net, reading a newspaper, watching TV. We have to filter information each hour of the day. It would be very useful if we only obtained the advertisements, tailored to our personal preferences. This personalisation is realised by taking into account available context parameters. This process is called narrowcasting which is unlike broadcasting bound to time, location and the audience interests. The most common type of narrowcasting are billboards that appear at semi-public places such as shops, bars, etc. In this paper, we present the description of an architecture for management of context-aware narrowcasting. We have developed a platform for displaying different types of media depending on the interests of the audience. The layout of each narrowcasting screen can be adjusted separately. Components are detailed for two scenarios: a billboard and an iDTV scenario.

Categories and Subject Descriptors

D.2.2 [**Design Tools and Techniques**]: Design Tools and Techniques; D.2.11 [**Software Architectures**]: Software Architectures; C.2.4 [**Distributed Systems**]: Distributed Systems; H.5.1 [**Multimedia Information Systems**]: Multimedia Information Systems; H.5.2 [**User Interfaces**]: User Interfaces

General Terms

Design, Human Factors, Management

Keywords

Personalisation, narrowcasting, profiling, iDTV, networks, advertising, context-awareness

1. INTRODUCTION

Because users are overwhelmed by a huge amount of content it becomes very difficult to identify content that matches their interests. Advertising concepts today are typically characterized by the *'one size fits all'* approach, where providing content is based on location and sometimes time. In order to display content which is appealing for the consumer we need to know the consumer preferences. Building up profiles of users and displaying content that matches these profiles is of prime importance.

Context-aware narrowcasting systems are the new way of advertising. They allow direct contact with the consumer by displaying advertisements only to those who are interested in a particular product. This is done by bounding the displayed content to time, location and the audience interests. Context can be collected from various sources like online social networks, webblogs, search queries, etc. Content can be specifically tailored to the users' interests by using techniques like ontologies, reasoning and rule engines.

In this paper, we present the design of an architecture for management of context-aware narrowcasting. Together with the operation of the platform we discuss the design in detail.

In the next section related work is discussed. Section 3 describes the modular architecture of the platform, the context management strategies and the layout management. Two use cases; one for a billboard deployment and one for an iDTV deployment are presented in section 4. Finally, future work and conclusion are mentioned in section 5.

2. RELATED WORK

Related to our research is the CASP project [1][2]. CASP is a context aware platform taking care of the aggregation and abstraction of context information for context-aware office and city services. The framework has a modular design which allows to include new architectual components plug-ins. SmartRotuaari [3] is an environment for context-aware mobile services which includes personalized advertisements. PersonalTV [4] is a tag cloud recommendation system for user-generated content, such as YouTube movies, which ex-

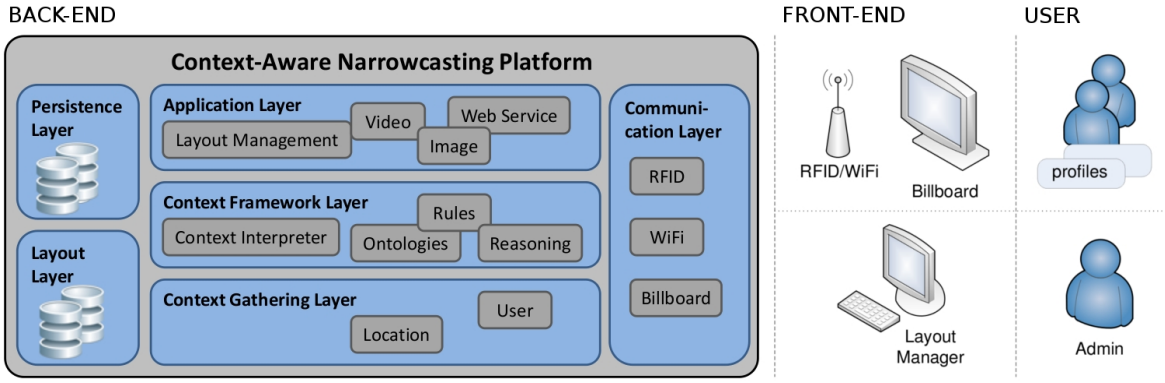


Figure 1: Overall Architecture Representation.

ploits social network relations. A content-based algorithm is used to predict the user's rating for unseen content based on his/her viewing and rating history. To the authors' knowledge, no previous work was identified on the design of an architecture for the management of context-aware narrowcasting.

In [5] an OSGi-based infrastructure manages context-aware services in a smart home environment is presented. The middleware is integrated with the OSGi service platform to ensure reliable and secure delivery of context-aware services.

3. ARCHITECTURE DETAILS

Figure 1 presents an overview of the architecture of the platform. Different layers can be discerned. All the layers are deployed in the back-end. The layout manager, where the administrator can adapt the layout of the billboard, and the billboard with RFID/WiFi interfaces are front-end devices. They connect to the back-end through the communication layer.

3.1 Layered Design

- **Persistence Layer.** The persistence layer takes care of the context information. This layer stores the content that is displayed on the billboard together with the context of the content. It also keeps track of all the profile information. This is information obtained from the context gathering layer. This is explained in more detail in section 3.2.

- **Layout Layer.** The layout layer stores the defined layouts. We support multiple layouts for different billboards, industries, etc. These layouts are made with a layout manager which is detailed in section 3.3.

- **Communication Layer.** The communication layer consists of all the devices that deliver context information to the platform. For example, a WiFi hot-spot that measures the signal strength from the connected devices. This information will be sent to the context gathering layer. Depending on the signal strength a user will have an impact on the content which is shown on the billboard.

- **Context Gathering Layer.** The context gathering layer is responsible for acquiring context information like location, music information, hobbies, etc from different sources.

- **Context Framework Layer.** The context framework layer retrieves the profiles from the audience and makes a pool of content built up by reasoning, ontologies and rules.

- **Application Layer.** The application layer uses the contextual information provided by the context gathering layer and processed by the context framework to provide content to the billboards. In order to achieve platform independence we opted for a web service. The web service take care of the communication between the back-end and the billboard.

We implemented the platform on an OSGi component system. OSGi technology is a dynamic module system for Java. The technology provides standardized primitives that allow applications to be constructed from small, reusable and collaborative components. These components, OSGi bundles, can be composed into an application and deployed. In addition, OSGi defines some standard toolkits for logging, configuring, security, user administration, etc to simplify development. Each layer consists of several OSGi bundles.

Installing, starting, stopping and updating bundles in the application layer is done by the web service. OSGi can execute this on the fly from anywhere in the network and at any moment. During the startup of a billboard the layout management bundle provides layout information to the context interpreter. This layout information contains the number of movies and images that have to be shown on the billboard and the position of the media. The provided content is dependent on both the layout information and context as well as the video and image bundles. This is explained in section 3.2 and 3.3.

3.2 Context Management

Context can be defined as any information that is used to characterize the situation of an entity [6]. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and application themselves.

To provide context to the content different modeling techniques exists. For this specific architecture we use Resource Description Framework (RDF) as a markup language based model [7]. This meta-data model has a subject-predicate-object structure which is called a triple. The subject is the

content described, the object its value and the predicate will denote a relationship between them. There are multiple triples for each content stored in an RDF database called relational graphs, each with a unique ID.

An Openlink Virtuoso [8] database is used to store this context in triples. Virtuoso is a middleware and database engine with specific support for RDF. Other database systems such as D2RQ [9], SPASQL [10] and DBLP [11] differ in the fact that Virtuoso combines mapping of relational data into RDF with native triple storage.

Ontologies define the relationship between objects. The Virtuoso database differs from a normal database because it does not only contain facts but also rules, defined in logical formulas. From these rules we can deduce new facts using automatic reasoning and model relations between concepts. We use ontologies to select content depending on the interests of the audience. We define our own standard ontology which is reusable in many different scenarios. The platform uses Jena Semantic Web Framework [12] for RDF.

3.3 Layout Management

In order to easily configure which layout each billboard should use, a layout manager was implemented. With a simple drag-and-drop system the advertiser can adjust the size of each canvas. He can define what kind of media will be played in each canvas and defines whether the media should be static or dynamic. Static media can be a banner with the name of a shop or a news channel. Dynamic media, mostly advertisements, are subjected to time, location and audience. The advertiser defines the screen resolution and refresh rate. For each advertisement area he can define whether to maintain the resolution of the displayed media or not. When all options are configured this layout is saved in a mysql database.

Figure 2: Graphical interface of the client.

The web service is used to play the content. We opted for a web-based design so no local installation is needed. The billboard only needs to access the browser. The web service is based on AJAX technology using GWT, Google Web Toolkit, which is a development toolkit for building complex browser-based applications, supported by most current browsers. GWT in combination with a Quicktime-plugin makes it possible to play movies on the billboard.

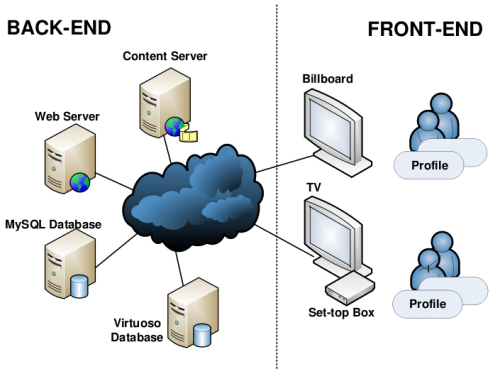


Figure 3: Billboard and iDTV Scenario.

4. SCENARIO DESCRIPTION

In this section, two use case scenarios will be discussed. In the first scenario our architecture will provide narrowcasting services for a billboard front-end while the second scenario elaborates on the use of our architecture with an interactive digital TV (iDTV) front-end, as shown in figure 3.

4.1 Billboard Scenario

The following scenario shows a billboard deployment in a crowded public environment. The billboard identifies, using RFID tags and readers, the presence of individuals and retrieves their profiles.

During the startup of the billboard, detailed in 3.3, a request for a screen layout is sent to the mysql database, as shown in figure 4. The Web Server will intercept this layout and pass it on to the billboard which is making a content request to the web server. The context gathering layer, as shown in figure 1, sends an identification request to the RFID sensors in the device layer. The identification information of recognized persons will be sent back to the context gathering layer. The layer now requests the profiles of these people at the Virtuoso database. The context gathering layer provides the collected information to the context framework layer.

The context framework layer uses the profiles and the context of the content, movies and images. Using ontologies, reasoning and rules, the system calculates the pool of content which matches the profiles of the audience and pass it on to the billboard using the web server as bypass. The billboard now makes a connection with the content server to collect the required content.

4.2 iDTV Scenario

In this section we discuss the implementation of our platform in an iDTV environment. We propose to implement our platform as an iDTV service using the Multimedia Home Platform (MHP) [13] of the Digital Video Broadcasting (DVB) consortium which is widely accepted in Europe. MHP enables the reception and execution of interactive, Java-based applications on iDTV.

We propose the same distributed architecture as for the billboard scenario, the STB runs a MHP application: the web service application. The iDTV provider takes care of the back-end. We combine the personalization of the advertise-

ments with displaying a television channel. This incorporates that for each home user the system needs to built up a profile and know when they are watching. Contextual information is gathered from the program viewing history and periodically sent to the back-end.

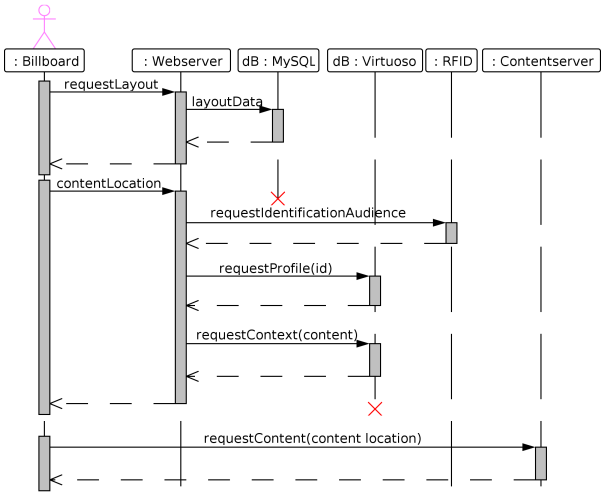


Figure 4: Sequence Diagram of Billboard Scenario.

The back-end provides a static screen layout to the platform. The STB identifies the people who are watching by their RFID tags and sends this to the back-end. The STB will receive the pool of content which matches the profiles and makes a connection with the server to retrieve the corresponding data. The downside to this scenario is that for each STB a separate datastream must be created. The operator has to make sure sufficient network resources are available.

This setup could also be used in a commercial setting. Using a web interface the customer designs his own screen layout with the screen manager, uploads and selects the advertisement that needs to be displayed based on their user profile. All this information will be sent to the back-end.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a management platform for context-aware narrowcasting. By using ontologies, reasoning and rules, the platform provides personalized media to the audience depending on the profile information of the users. The Virtuoso database, which provides RDF support, is used to store profile and content contextual information. It is shown that the platform supports several scenarios: billboard and iDTV sceniario. The layout management tool creates dynamic layouts which are loaded by the billboard. The platform was built compliant to OSGi standards which makes it easy to manage and extend its functionality.

Instead of using RFID technology we are currently developing an adaptive system with real-time face detection and recognition [14]. With eye tracking we are able to determine which people are watching and how long they are watching specific content. This way we can add the context information of that media to the person's profile with a weight in function of the watched time.

6. ACKNOWLEDGMENTS

This work is partly funded by the IBBT SBO GR@SP project. Tim Wauters acknowledges the Fund for Scientific Research - Flanders (F.W.O.-Vlaanderen) for their support through a postdoctoral fellowship.

7. REFERENCES

[1] Strobbe M., Hollez J., De Jans G., Van Laere O., Nelis J., De Turck F., Dhoedt B., Demeester P., Janssens N., Pollet T., *Design of CASP: an open enabling platform for context aware office and city services*, Proceedings of the 4th International Workshop on Managing Ubiquitous Communications and Services, 2007, pp. 123-142.

[2] Van Laere O., Strobbe M., Dauwe S., Dhoedt B., De Turck F., Demeester P., Verde O., Hulsken F., *Interest Based Selection of User Generated Content for Rich Multimedia Services*, Proceedings of the 9the International Workshop on Image Analysis for Multimedia Interactive Services, 2008, pp.36-40.

[3] Ojala T., Korhonen J., Aittola M., Ollila M., Koivumäki T., Tähtinen J., Karjaluoto H.,*SmartRotuaari - Context-aware Mobile Multimedia Services*, Proceedings of 2nd International Conference on Mobile and Ubiquitous Multimedia, 2003, pp. 9-18.

[4] De Pessemier T., Deryckere T., Martens L., *Context aware recommendations for user-generated content on a social network site*, Proceedings of the seventh european conference on European interactive television conference, 2009, pp. 113-136.

[5] Gu T., Pung H. K., Zhang D. Q., *Toward an OSGi-based infrastructure for context-aware applications*, IEEE journal on Pervasive Computing, 2004, pp. 66-74.

[6] Dey A. K., *Providing Architectural Support for Building Context-Aware Applications*, PhD thesis, College of Computing, Georgia Institute of Technology, December 2000.

[7] Beckett D., *Rdf/xml syntax specification (revised)*, http://www.w3.org/TR/rdf-syntax-grammar/, 2004.

[8] Erling O., Mikhailov I., *RDF Support in the Virtuoso DBMS*, Proceedings of the 1st Conference on Social Semantic Web, 2007, pp. 59-68.

[9] Bizer C., Seaborne A., *D2RQ - Treating Non-RDF Databases as Virtual RDF Graphs*,Proceedings of the 3th International Semantic Web Conference, 2004.

[10] Prudhommeaux E., W3C, *SPASQL: SPARQL Support In MySQL*, http://xtech06.usefulinc.com/schedule/paper/156

[11] Chen H., Wang Y., Wang H., *Towards a Semantic Web of Relational Databases: a Practical Semantic Toolkit and an In-Use Case from Traditional Chinese Medicine*, Proceedings of the 5th International Semantic Web Conference, 2006, pp. 750-763.

[12] Jena - A Semantic Web Frame *Context Framework Layer*work for Java, *http://jena.sourceforge.net*, 2009

[13] Multimedia Home Platform, *http://www.mhp.org*, 2009.

[14] Viola, P., Jones, M.J., *Robust real-time face detection*, Int. J. Comput. Vision 57(2), 2004, pp. 137-154.

Integrating Interactive TV with the Web

Cristina Carmona
CITIC
6 Marie Curie, PTA, 29590
Málaga (Spain)
Phone: +34 952028610
ccarmona@citic.es

José A. Rodríguez
CITIC
6 Marie Curie, PTA, 29590
Málaga (Spain)
Phone: +34 952028610
jarodriguez@citic.es

José A. Hurtado
CITIC
6 Marie Curie, PTA, 29590
Málaga (Spain)
Phone: +34 952028610
jahurtado@citic.es

J. R. Salinas
CITIC
6 Marie Curie, PTA, 29590
Málaga (Spain)
Phone: +34 952028610
jsalinas@citic.es

ABSTRACT
Nowadays, there are two facts that are absolutely positive. On the one hand, the web is currently the biggest information repository available. On the other, almost everybody have a TV. That means that most web users are also TV users. In the last years, much research work regarding digital TV has been made. Among many other things, mechanisms to execute interactive applications in a digital TV decoder have been established. This paper presents a framework aimed to integrate web content into a TV application easily. This approach proposes the use of some tools able to extract the relevant data from the web page in order to be used by TV applications. This framework has been successfully tested with the development of the MHPTur Project, which is a TV application for tourism information based in data provided by the official tourism website of Andalucía.

Categories and Subject Descriptors
D.3.3 [**Programming Languages**]: Language Contructs and Features – *frameworks, modules, structures.*

General Terms
Performance, Design, Reliability.

Keywords
Interactive Applications Framework, Multimedia Home Platform, Web.

1. INTRODUCTION
Interactive TV enhances the conventional TV services allowing user interaction. In the last years, there have been many works in this area. Part of this work is focused in defining standards to establish how to work with digital television and its applications. Multimedia Home Platform (MHP) [1]is a standard designed by the DVB Project and defines a generic interface between interactive digital applications and the terminals on which those applications execute. Most countries in Europe have chosen MHP to develop their TV applications. Another approach is Open Cable Application Platform (OCAP) [2], that is the software environment standard developed by the US cable TV industry. It is based in the DVB proposal, but some specifications that are not used in the cable environment are removed and replaced by others with the equivalent functionality.

Other works are oriented to TV applications development. In 2002 Nionex presented Pontegra [3], the first MHP browser based in the open standard DVB-HTML, which makes the development of interactive TV services as easy as programming HTML-Web pages. Since then, many other approaches to MHP browsers were presented relying on XML descriptions to build applications: tmPlayer from Tmira [4], Yambo from Cineca [5], Andago T-Browser from OpenTDT [6] or the Browser XML [7] developed by CITIC. Several research fields try to use existing procedures and technologies with TV applications. For example, the use of interactive TV with recommender systems to help users to find contents appealing to them among all the irrelevant information [8] [9]; with social networks, developing new user's experiences based on the content shared by others [10]; etc.

Most of these applications make an intensive use of the return channel to exchange information with the "main server", usually a web server. In this paper, a framework to easily integrate web content into a TV application is proposed. The idea is to have only one interactive TV application broadcasted, the Browser XML, and let the web do all the process. When the user makes a query, the information is sent to the web, which produces a response. This response has to be translated to an appropriate XML format and finally be shown to the user through the Browser XML. Different modules within the framework are responsible for all this processing that must be transparent for the user and must work for all the websites.

To evaluate this system, an interactive TV application for the tourism sector has been designed. The MHPTur Project has been granted by the *Conserjería de Comercio, Turismo y Deporte - Junta de Andalucía* [11] (Ref.: 38/08).

Following section is focused in describing the framework design. Next, the MHPTur project is presented, as an evaluation of the framework to test its usability and performance. Finally, conclusions show the evaluation results and future work.

2. FRAMEWORK DESCRIPTION
As briefly explained in previous section, this framework is an intermediary system between a web application and a TV application. Figure 1 shows the basic system architecture. Besides the audio and video, a Set-Top-Box (STB) can receive applications, like the Browser XML that is an MHP application that translates XML documents into MHP screens, defining its appearance and behavior. So, it is possible to build applications dynamically and almost automatically. It is only necessary to design some tools to build those XML documents from the content in the web: a web scrapper (*Jet-Tag*) and a web content syndication tool (*MHP-Feeder*).

Web scrapping is a technique in which some program extracts data from a web page in order to be used by another program. Thus, Jet-Tag is able to obtain all the information from a web page, providing semantic information about the page structure. JetTag builds a template for the web page, that is, a definition of

the page layout identifying where the relevant information is located. Then, it will use this template together with the web page to obtain the data automatically.

MHP-Feeder is a program that has a list of specific feeders and can be scheduled to check for new content periodically. When ordered, MHP-Feeder asks all the feeders in its list to check if they have new content. If so, this content is downloaded. Afterwards, Jet-Tag templates are applied to obtain relevant information from these documents. Finally, this data is given the appropriate format for the Browser XML and the carousel is updated. This update is made periodically and in an autonomous way.

However, not all the information in the web can be broadcasted this way. For instance, if the application has to make some queries to a database to get some data, it is impossible to include the whole database in the carousel. In this case the user has to fill in a form to get the filtered data and the STB must communicate with a proxy web, using the return channel.

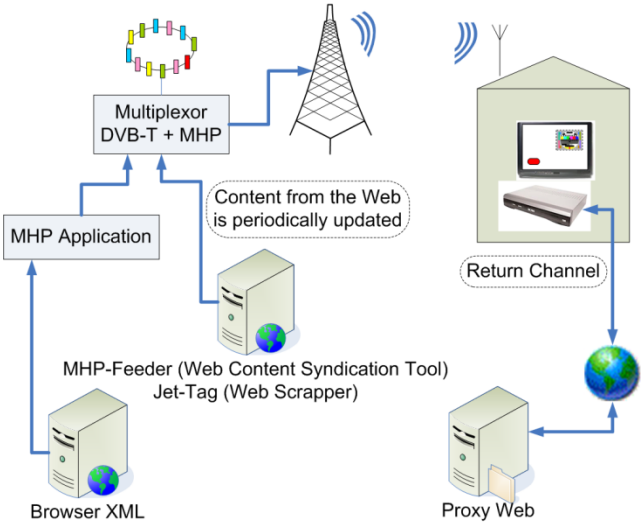


Figure 1. Framework Architecture

Summarizing, with this model only one application must be broadcasted, that is, the Browser XML. The browser shows screens or applications described by XML documents generated dynamically from the web content. For those web pages that are only informative (no user interaction), the MHP-feeder tool builds the XML files and update the carousel periodically. If the web pages need user interaction, it is necessary to use the return channel to connect with the proxy web to resolve the query. The proxy web has to translate the web response into an appropriate XML document to be shown in the Browser XML. These tools are described in detail in next subsections.

2.1 Browser XML

The Browser XML is a standard MHP application, which makes it possible to run a complete interactive application, written entirely in XML format with the possibility of adding new feature extensions through plug-ins written in Java.

The Browser XML provides a higher level of abstraction for application development making it easy to write several tasks, which are more tedious to write in Java. The largest part of the code needed to perform typical operations with the MHP API is placed in the browser code, which is shared with all applications, thereby saving space in the object carousel and reducing the startup time. More information about the Browser XML can be found in [7].

2.2 Jet-Tag

HTML tags in a web page define its structure and presentation, but do not give any information about its semantic. Jet-Tag is able to analyze the page structure to provide some semantic information about the relevant tags within that page for the user.

Jet-Tag is a web scrapper based on the vectorization of HTML documents. For each HTML tag, Jet-Tag builds a vector with a set of values representing this tag, so those vectors can be used to find the distance between two tags in the document. Given a HTML tag within the document as a prototype, Jet-Tag is able to find all the HTML tags similar to the prototype. Similarity is defined using the distance between a tag and the prototype tag, so those elements whose distance is less than a user defined threshold are selected. As a result, Jet-Tag produces a XML document with one element for each selected tag. Jet-Tag configuration for that HTML document can be saved as a template, so it can be applied again to the document to get up to date data.

```
<extraction>
 <record>
  <date>From 19/03/2009 To 13/12/2009</date>
  <place> Cádiz, Granada, Málaga</place>
  <image>
   <jettag_image type="jpeg">
    ffd8ffe000104a46494600010100000100010000 ...
   </jettag_image>
  </image>
  <type>Concert</type>
  <name>Pasión Vega</name>
 </record>
.........
<extraction>
```

Figure 2. Jet-Tag: XML file

The whole process can be easily done using a web application and is detailed as follows:

1. Prototype Selection: once a HTML document has been loaded, the user must select the element to extract. This element will act as the prototype, so similar elements within the document will be extracted.
2. Similarity Threshold: the user must select the threshold that defines similar elements. If a threshold is 0, only the prototype element will be chosen, as the threshold increase, more elements will be selected.
3. Field Definition: for the prototype element, the user has to select the fields of interest. A name must be given for each selected field within the element, so this name will be the element name for that field in the final XML file. It is possible to save this description as a template, so it can be

used subsequently with the same HTML document. This template represents the relevant semantic information of the web page for the user.

4. Extracted Data: finally, the template is used with the HTML document to extract the information and it is shown in a XML format. There is one record for each tag in the HTML document which is similar to the prototype one. Since the template is saved, it can be used later with the same HTML document to get a new XML file with up to date information. An example of an XML file is shown in Figure 2.

2.3 MHP-Feeder

MHP-Feeder is a web content syndication tool and is used to find new information to be added in the TV application carousel, so it can be accessed without using the STB return channel. Usually, this kind of information changes periodically and with no user interaction.

The process for generating Browser XML files using MHP-Feeder is described in Figure 3. The information gathered by MHP-Feeder is available on the web, so the first step is to use Jet-Tag with the URL of the HTML documents relevant for the application in order to obtain templates describing them. For each template, MHP-Feeder has a specific module which is responsible to apply its template to the URL of the document, getting the current information from the web. MHP-Feeder collects all the XML documents from each module and builds the final XML document in the Browser XML format. MHP-Feeder is executed periodically and in an autonomous way.

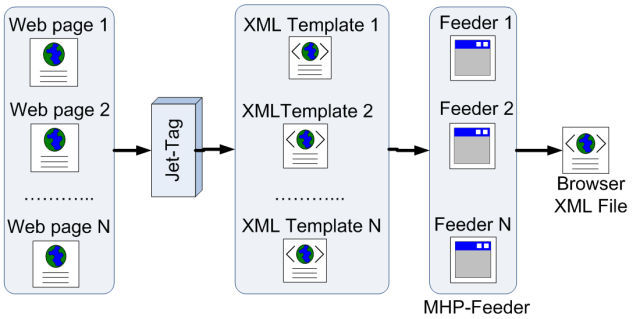


Figure 3. Generating Browser XML Files with MHP-Feeder

2.4 The Proxy Web

The proxy web has to receive the users' petitions from the STB, process them and, finally, build a response to be sent to the STB. Figure 4 shows the architecture of the proxy web.

An application may have many forms for different requests. Each request is received by the queries module, which has to send it to the specific sub-module. This sub-module has to make the web request, get the result and translate it to a Browser XML format. This architecture allows making this proxy web extensible and domain independent, that is, sub-modules can be added or removed at any moment, with a very little change in the main queries module.

The sub-module XML file is returned to the queries module and, finally, sent to the MHP decoder. If the web response contains some data that cannot be included within the XML file, like image files, this resources must be stored in a content repository in the proxy web, so in the XML file, the resource can be referenced with a URL. This content repository is managed by the dynamic content module.

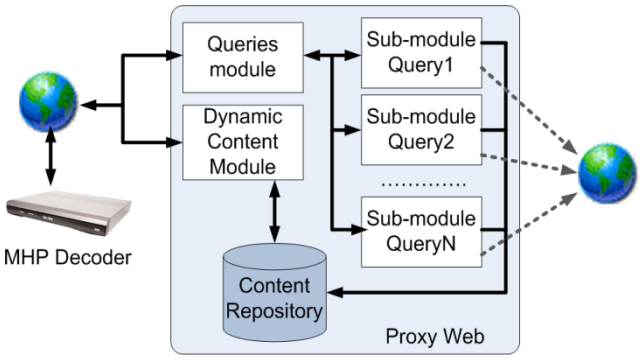


Figure 4. Proxy Web Architecture

In order to improve the return channel performance, some actions have been taken: caching and proactive requesting. Whenever a search is done, the result is cached, so if this query is repeated there won't be necessary to perform a web request again. Also, when the search results are cached, a proactive requesting for the details of each element in the list is performed, so if the user asks for the detail, it will be already cached.

3. MHPTur

An MHP application for the tourism sector in Andalucía has been designed using this framework. Different websites with information about Andalucía have been used as the source of information for this project (weather, festivals, hotels, restaurants, etc). The relevant information from these websites must be classified depending on the user interaction. As a result, there are three groups of web pages:

1. Static pages: this group includes pages with general information about the main towns, geographical location, festivals, etc.
2. Dynamic with no user interaction: this group includes pages with the information about events, news and weather. For this project, this information is updated every hour.
3. Dynamic with user interaction: this group includes pages related with searching facilities for hotels, restaurants, golf courses and sightseeing.

Although all this information is available on the web and it could be accessed via the return channel, it is recommended to use it only when it is totally necessary, so most of the information must go in the carousel with the application. This way, most of the information is available for everyone who has a STB, even if it does not have a return channel.

For the first group, static Browser XML files have been designed. For the second group, three feeders for MHP-Feeder have been developed. All of them have the same structure and behavior, but the information they manage is different. Finally, for the third group, four sub-modules for specific queries have been implemented in the proxy web. The main queries module only needs a list of active sub-modules in order to know where to send each request. As in the previous group, all these sub-modules have the same behavior but deal with different kind of data.

Some screenshots of the MHPTur application are shown in Figure 5, specifically the hotel search facility. First, the user has to fill in a form to set up his preferences (1). Afterwards, a search in the website is performed and the result is a list of hotels that match the preferences established by the user. This list is formatted in a suitable way for the Browser XML and shown to the user (2). Then, when the user selects a hotel to see the detailed information, a new request is made (3).

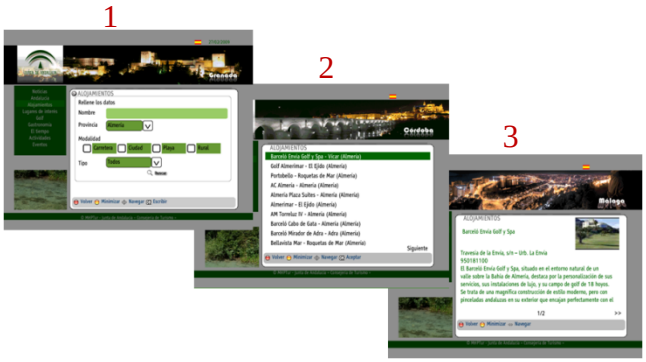


Figure 5. MHPTur Application

4. CONCLUSIONS

This paper proposes a framework to make easy the integration of web content into TV applications. Some tools to extract relevant information from HTML pages and to perform the communication using the return channel have been designed.

The MHPTur Project has been developed for testing purposes. Regarding usability, from the developer viewpoint, it has been very easy and fast given that once a feeder or sub-module is implemented, the rest of the same type are done with very few changes. That makes this project extensible with low effort. Regarding performance, some factors have been taken into account to improve the framework. The first one is the website processing time. The delay produced by the website chosen as a source of information for the Proxy Web is out of the proxy scope of control. Some mechanisms of caching and prefetching have been implemented in order to reduce the impact of this delay in the framework. Another factor that influences the performance is the carousel loading bitrates. Increasing the number of applications on the carousel implies a bigger amount of broadcasted data and the loading time increase, as bitrates is fixed by service providers. As demonstrated in [7], with the use of the Browser XML, the framework achieves better time responses while minimizing the applications size.

The development of the MHPTur Project shows that this framework makes easy to provide this kind of services with the benefits they entail, like user-friendliness (a very simple user interface), information available for non-expert users (bridging the digital divide), web information accessible through TV, etc.

5. ACKNOWLEDGEMENTS

This work has been supported by Centro Andaluz de Innovación y Tecnologías de la Información y las Comunicaciones (CITIC), as part of its Research and Development strategic line in Interactive Digital Terrestrial Television. CITIC, as Research and Innovation Center in the Information Technology (IT) and Communications areas, works by developing high technology projects in these fields, being its main aims the IT and Communication innovation, Information Society advance contribution and research results' diffusion and transferring promotion.

6. REFERENCES

[1] Digital Video Broadcasting (DVB) - Multimedia Home Platform (MHP) Specification 1.2.2 (June 2009): http://www.mhp.org/specs/a140_tm4242_MHP_122_refactored.pdf

[2] OCAP website: http://www.opencable.com

[3] Pontegra by Nionex: http://www.medianetvlaanderen.be/content/user/File/verslagen/pontegraShort.pdf

[4] Tmira Solutions website: http://www.tmira.com

[5] Cineca website: http://www.cineca.tv

[6] OpenTDT website: http://www.opentdt.com

[7] Hurtado, J.A., Ruiz, B. Padilla, R. and Salinas, J.R. 2009. "An Optimized Framework for Interactive TV Services Broadcasting". In adjunct proceedings of the 7th European Interactive Television Conference (Leuven, Belgium, June 2009).

[8] Blanco, Y., Pazos, J.J., Gil, A., Ramos, M., López, M., García J., Fernández, A., Díaz R.P. and Bermejo, J. 2008. "An MHP framework to provide intelligent personalized recommendations about digital TV contents". Software: Practice & Experience. Volume 38, Issue 9 (2008), pp. 925-960.

[9] Zaletelj, J. 2008. "Recommender System for the Multi-channel TV Production". In proceedings of 6th European Interactive Television Conference (Salzburg, Austria, July 2008).

[10] Alliez. D. 2008. "Adapt TV paradigms to UGC by importing social networks". In adjunct proceedings of the 6th European Interactive Television Conference (Salzburg, Austria, July 2008).

[11] Junta de Andalucía, Consejería de Turismo, Comercio y Deporte website: http://www.juntadeandalucia.es/turismocomercioydeporte

A study on location of context-aware services for Mobile Digital Television

André Valdestilhas
Address: América do sul, 812
CEP: 09270-410
Santo André – São Paulo - Brazil
Tel: 55-11-8795-9880
firmao@yahoo.com.br

ABSTRACT
The popularity of gadgets such as cell phone and GPS (Global Position System) for cars, or both in one device, with the use of digital television is increasingly pervasive in our society in this respect the article is to discuss the availability of content the shape of local services for these devices by locating in an area defined by resource use of GPS and triangulation techniques already known for use in electronic maps, and is now suggested to be applied to digital mobile TV. Thus, it is possible to provide a range of local services such as proximity to hospitals, pharmacies, theaters and local advertisements or send the coordinates of your location to the content provider, to request any special service, such as delivery of a product or charging for services. Seeking validation, we propose an architecture for providing these services by content provider.

Categories and Subject Descriptors
K.m [MISCELLANEOUS]

Keywords
Mobile Digital TV, GPS, Content Distribution.

1. INTRODUCTION
Electronic devices such as mobile phones are evolving where cameras, GPS, Bluetooth, Internet access and digital television are already common place in these devices.

The aim of this paper is to propose an approach to provide context sensitivity for application in mobile digital television, using concepts obtained in [3], used in conjunction with the Ginga-NCL [1] and concepts of location of mobile devices presented in [2].

2. CONTEXT AWARE COMPUTING
Context aware computing is a term that originates from the area of ubiquitous computing, which was first presented by Mark Weiser in 1991 [6].

According to Dey [4] "any information that can be used to characterize the situation of entities (be it a person, place or object) that is relevant to the interactions between users and applications," which provides a definition for context computing that, although generic, is rather referenced in the literature.

In Rocha [5] the main purpose of computing sensitive or aware of the context is to allow dynamic adaptation of applications and services for different purposes, such as proper use of resources and creating more User interfaces with rich and effective.

Categorization of context:

Studies were made [7] [8] and [9] in order to categorize contexts, which can be presented as:

- **Perception of context**: is the ability to detect the context and present the information that describe the User.
- **Association of contextual information**: to associate contextual information data. For example, in a meeting, the persons present and the meeting venue can be associated with the minutes of the meeting.
- **Resource Discovery**: allows applications to discover and explore resources and services relevant to a particular context.
- **Adaptation context**: this feature is used in the literature to describe cases where the context causes an action and cases where the context is used to modify or adapt services.

3. RELATED WORK
A These related works presented here are intended to discuss the integration between them in order to propose a better solution for contextualizing content for Mobile Digital Television.

According to studies done in [1] which presents the possibility of the use of middleware Ginga Brazilian resources of Mobile Digital TV due to the Ginga-NCL.

Figure 3 illustrates the modular architecture of Ginga-NCL for portable devices, which is divided into two logical subsystems: the machine Presentation Ginga-NCL and Ginga Center.

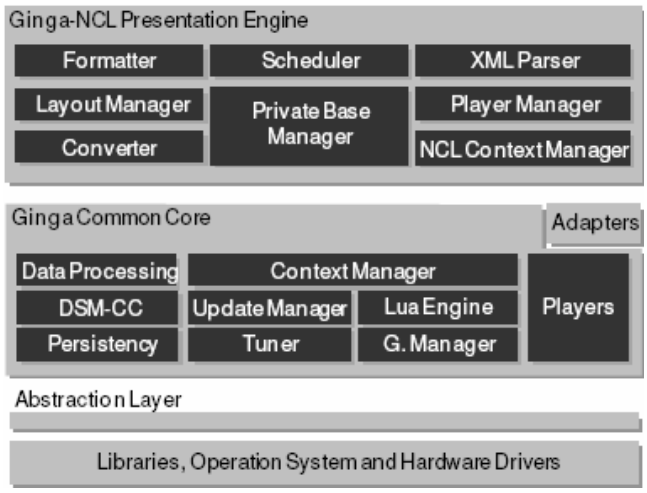


figure 1. Modular architecture of Ginga-NCL.

- The machine Presentation Ginga-NCL is a logical subsystem can initiate and control applications in NCL are the converter, the formatter.
- The Center Ginga is responsible for providing services to the Machine Presentation Ginga.

Work available in [2] called PlaceLab presents an architecture which enables the use of resources Wi-Fi and Bluetooth. The architecture consists of three main elements, as shown in Figure 2:

- Signs of transmission in the environment;
- Database, which stores information on these signals and their locations, and;
- The system that runs on the computing environment of the User using this information to estimate your current position.

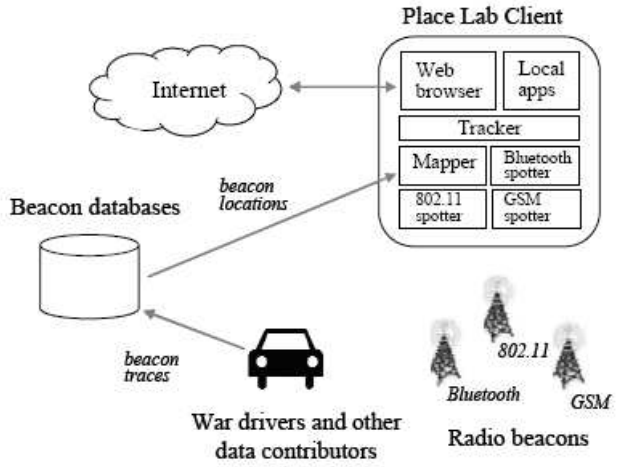


figure 2. PlaceLab Architecture.

The contexTV [3] is a proposal for an architecture like Client-Server in which the client deals with features of mobile devices like cellular handsets and the server, which can be located on a content provider of DTV, Figure 3 shows the proposed architecture the contexTV.

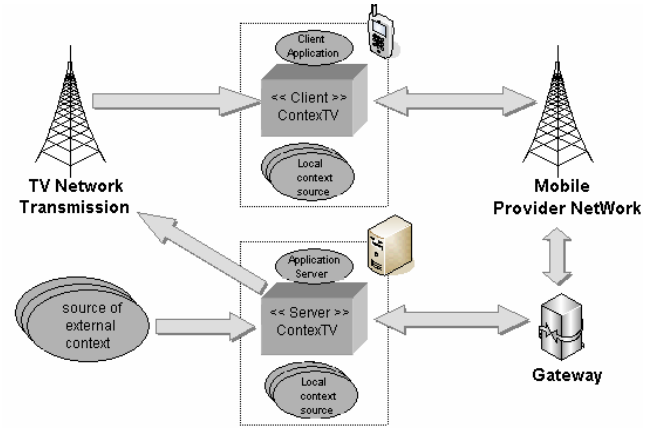


figure 3. ContexTV Architecture.

One reason for the use of the approach presented in contexTV [3] is also a need presented in [1] which addresses the problem of having some sort of connectivity with wireless networks (IEEE 802.11), which combined with the approach presented in [2] can complete this pending.

4. THE PROPOSED ARCHITECTURE

The architecture shown in Figure 4 is to propose a way to provide the Mobile Digital Television addressed by Ginga [1], by aggregating resources for locating mobile devices available in [2] and contextualization of the content transmitted by the provider, discussed in [3].
Thus it is possible the deployment of Digital Television Mobile context-sensitive in Brazil and the technology used varies according to region, for example, Wi-fi could be used where the resource was available, if not only the content transmitted via broadcast site.

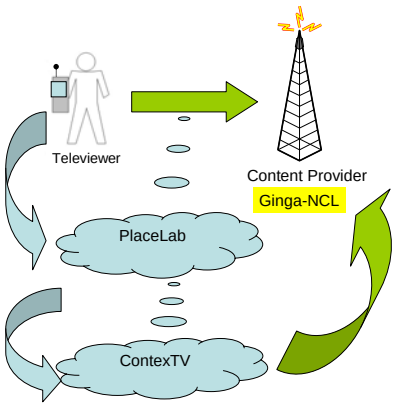


figure 4. Proposal Integration Architecture.

4.1 Description of architectural elements

The core Ginga, the tuner module is responsible for receiving digital TV content for handheld devices by broadcast content providers. The interactive applications may arrive at the portable device multiplexed in content, or another network interface (eg, data channel offered by mobile service provider, Bluetooth or IEEE 802.11, among others.) [1].

This makes possible the use of Place-Lab [2] together with the Ginga, because it is based on signal transmissions through blueetooth and Wireless (IEEE 802.11).

The ContexTV [3] would act in a way to capture and process data stored by the Place Lab-and so enable the sensitivity to context.

5. CLOSING REMARKS

This work differs from other works because it covers the use of middleware Ginga Brazilian [1] and the use of concepts obtained ContexTV [3], indicating a possible integration of related work.

Because it is based on concepts that are being used successfully in the scenario of the Brazilian Digital TV, it is believed that this approach to development will be able to provide a way to provide the Brazilian mobile TV capacity of context-awareness and help the viewer's location these services.

6. REFERENCES

[1] Cruz, V. M. *Ginga-NCL para Dispositivos Portáteis*. Masters dissertation; PUC-Rio, DI, 2008.

[2] LAMARCA, A. *Self-Mapping in 802.11 Location Systems*. In Proceedings of Ubicomp 2005. Tokyo, Japan. set. 2005.

[3] ANDRADE NETO, F. C.; FERRAZ, C. Arquitetura de suporte a aplicações de tv digital interativa sensíveis a contexto. In: II Workshop de TV Digital - SBRC 2006.

[4] Dey, A.K. et al, “A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications”, HCI Journal, 16(2-4), 2001, 97-166.

[5] da Rocha, R. C. and Endler, M. 2005. Evolutionary and efficient context management in heterogeneous environments. In Proceedings of the 3rd international Workshop on Middleware For Pervasive and Ad-Hoc Computing (Grenoble, France, November 28 - December 02, 2005). MPAC '05, vol. 115. ACM Press, New York, NY, 1-7.

[6] WEISER, M. The computer for the 21st century. Scientific American, v. 265, n. 3, p. 94-104.Disponível em: <http://www.ubiq.com/hypertext/weiser/SciAmDraft3.html>. Acesso em: 19 ago. 2009.

[7] SCHILIT, B. N.; ADAMS, N.; WANT, R. Context-Aware Computing Applications. In: IEEE WORKSHOP ON MOBILE COMPUTING SYSTEMS AND APPLICATIONS, 1., 1994, Santa Cruz. Proceedings. Los Alamitos: IEEE, 1994. p. 85-90.

[8] PASCOE, J. Adding Generic Contextual Capabilities to Wearable Computers. In: INTERNATIONAL SYMPOSIUM ON WEARABLE COMPUTERS, 2., 1998, Pittsburgh. Proceedings. Los Alamitos: IEEE, 1998. p. 92-99.

[9] DEY, A.K.; ABOWD, G.D. Towards a Better Understanding of Context and Context-Awareness. Atlanta, College of Computing, Georgia Institute of Technology, 1999. (Technical Report).

Automatic video search using semantic concepts

Mats Sjöberg
mats.sjoberg@tkk.fi

Ville Viitaniemi
ville.viitaniemi@tkk.fi

Markus Koskela
markus.koskela@tkk.fi

Jorma Laaksonen
jorma.laaksonen@tkk.fi

Adaptive Informatics Research Centre
Aalto University School of Science and Technology
P.O. Box 15400, FI-00076 Aalto, Finland

ABSTRACT

This paper describes automatic video search using semantic concept detection, and how it is implemented in our PicSOM system. The demand for such methods is growing because of the exploding growth of digital video data produced today, ranging from professionally produced TV programming to individuals sharing videos online. Content-based methods address the problem of finding relevant information in large amounts of visual data, which by nature are very hard to search and index automatically. The semantic gap between the high-level concepts understood by humans and the low-level information of computer systems is partially bridged by modelling mid-level semantic concepts. The performance of our system is judged in the video retrieval setting of the international TRECVID 2008 and 2009 evaluations, comparing favourably with the competing state-of-the-art systems.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing; I.2.6 [**Artificial Intelligence**]: Learning—*Concept learning*

General Terms

Algorithms, Experimentation, Performance

Keywords

video search, concept detection, support vector machine

1. INTRODUCTION

Digital video has become commonplace, both in professional use and in various consumer products such as camcorders, webcams, digital TV, mobile phones, video sharing websites, CCTV surveillance, and virtual and augmented reality applications. While the capturing, storing, and transmitting of digital video has steadily become easier and more cost-effective, the current methods for the automatic analysis and semantic representation of the video content itself are considerably less mature.

Content-based video retrieval addresses the problem of finding visual data relevant to the users' information needs from video repositories [4]. This is generally done with querying by examples and measuring the similarity of the database objects (keyframes, video clips) with *low-level features* automatically extracted from the objects. Generic low-level features are often, however, insufficient to discriminate content well on a conceptual level. This "semantic gap" is the fundamental problem in content-based multimedia retrieval.

The modelling of *mid-level semantic concepts* (events, objects, locations, people, etc.) attempts to fill, or at least reduce, the semantic gap. Indeed, in recent studies it has been observed that, despite the fact that the accuracy of the concept detectors is far from perfect, they can be useful in supporting *high-level indexing and querying* on multimedia data [1]. This is mainly because such semantic concept detectors can be trained off-line with computationally more demanding supervised learning algorithms and with considerably more positive and negative training examples than what are typically available at query time.

In this paper, we first describe briefly the relevant parts of our PicSOM multimedia analysis and retrieval framework [2], including video corpus preparation, concept detection and concept-based automatic search. We then use the large-scale experimental setup of the TRECVID evaluation campaign to relate the performance of our system to other systems in the video retrieval community for automatic search (i.e. without human input in the loop).

2. PARTS OF A VIDEO SEARCH SYSTEM

This section gives an overview of the components of a video retrieval system in general and the implementation in our PicSOM system in particular. For brevity, only a brief summary of the system is provided in this paper. More detailed descriptions can be found in our annual TRECVID workshop papers (e.g. [3]). The general architecture of the PicSOM video search system is shown in Figure 1.

The operation of a video retrieval system generally consists of two phases: an offline *preparation phase*, where a database is processed and indexed, and the online *search phase*, in which responses to queries must be generated in a reasonable time for human interaction.

In the preparing phase the video corpus is usually first segmented into shots and a number of low-level visual, audio and textual feature descriptors are extracted from each shot and content-based indices are prepared based on the features. Shot-wise detectors are trained on annotated data, either from the same corpus or they might be trained previously on similar data. The detectors apply supervised learning techniques to learn the mapping between low-level shot features and the annotation concepts. The preparation is allowed to be time-consuming as it is intended to be performed off-line prior to the on-line use of the retrieval system.

In the search phase, the system is set to respond to video

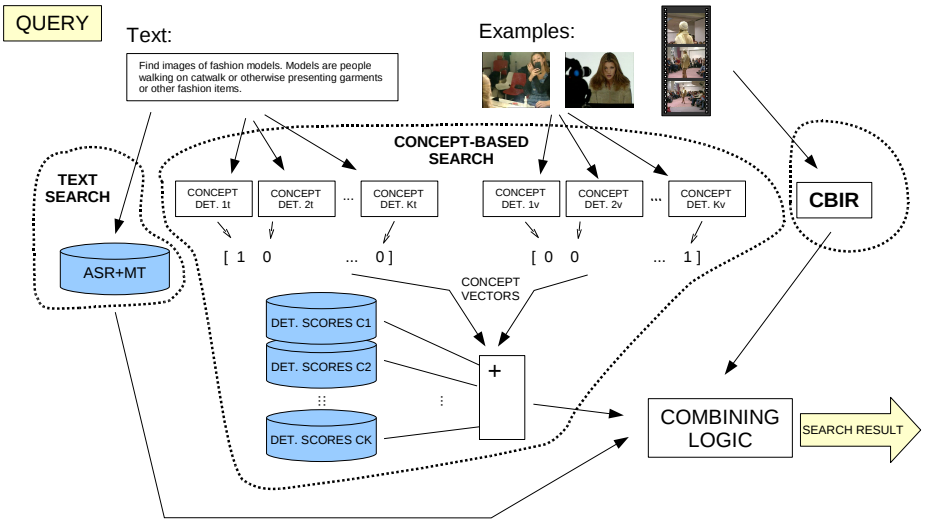


Figure 1: Overview of video search with PicSOM.

queries. The queries may constitute of textual phrases, image and video examples, or of any combination of them (see an example in the top part of Figure 1). The retrieval system can then employ some or all of the following search modalities: *text search*, *concept-based retrieval* and *content-based retrieval*. The retrieval result is a list of video shots, ranked in the order of decreasing predicted likelihood to match the query. The system operation in the search phase is intended to be sufficiently fast for interactive use.

2.1 Shot segmentation and keyframe selection

The first task of the preparing phase for a comprehensive video retrieval system is to segment the video corpus temporally into sequential basic units. Depending on the video material, such a segmentation can be performed on various levels, such as stories, events, scenes, groups, sequences, and shots. For scripted content, the basic semantic unit is the shot, as shots are intrinsically linked to the production of the video. As shots can usually be identified by automatic methods with a reasonable accuracy and they provide a suitable representation level for the higher-level video analysis tasks, the contemporary video retrieval systems customarily treat the shot as the basic unit of retrieval. In the TRECVID evaluations, a master definition of shots is provided by the organisers. In cases where the shot boundaries have not been available, we have applied one of two shot segmentation algorithms available in the PicSOM system [3].

To facilitate the usage of image-based retrieval, we also extract keyframe images from each video shot. If a video shot is short and contains only one visually homogeneous scene, a single well-chosen keyframe can compactly express the most central visual characteristics of that shot. The keyframes are also useful as still visual representations the shots when presenting a collage of several results to the user. In the PicSOM system the keyframe is selected by a heuristic method which tries to pick a "typical frame" (i.e. similar to the average of all frames) which is also close to the temporal centre of the shot, and that does not have rapid movement.The last requirement is to avoid motion blurring effects.

2.2 Low-level features

Using the video or image data directly in retrieval is typically not feasible due to the high dimensionality of the data. Extracted low-level features should thus ideally be of reasonable dimensionality and at the same time be discriminative of semantic differences in the data, i.e. the feature extractors should be sensitive to those characteristics of the raw data that are relevant to the human perception of the media contents. For video analysis in particular there is an opportunity to combine several data modalities, such as keyframe images, video motion, audio and text from speech recognition or subtitling. From these modalities diverse feature representations can be extracted to represent different relevant and complementary aspects of the underlying data.

Image features. The PicSOM system uses a wide range of image features that have been added over the years of its development. In addition to standardised MPEG-7 descriptors and some non-standard image features developed in-house the PicSOM system employs state-of-the-art bag-of-visual-words (BoV) features, in which images are represented by histograms of local image descriptors.The latter category includes in particular SIFT-based descriptors together with the Harris-Laplace interest point detector or dense sampling. These features are described further in [3].

Video features. In many cases the static visual properties of a single video frame are not enough to describe the salient features of the full scene. In some situations the motion of objects or the camera might be semantically significant, for example in distinguishing between a ball that is rolling from one that is still. Also, the dynamic properties may in some cases make the computational learning problem easier. For example it may be easy for a human to recognise a running person even from a still keyframe image, but such videos are surely easier to distinguish based on the temporal properties of the person moving across the scene. The set of video features used in PicSOM include temporal extensions to some of the still-image features, and specialized motion features [3].

Audio features. Most video shots include a sound track, containing for example music or different environment sounds.

A certain distinctive musical tune may perhaps indicate the beginning of a news broadcast, or indicate for example the occurrence of an action scene in a movie. A crowd cheering in a football game is a strong cue of an important event such as a goal being scored. In PicSOM we have used implementations of the popular mel-scaled cepstral coefficients feature (MFCC) as an audio feature.

2.3 Textual search

Often, the video material includes textual data or metadata that can facilitate text-based indexing and retrieval. Textual data for video shots may originate e.g. from speech recognition, closed captions, subtitles, or video OCR. The textual search module can easily be implemented as a separate component whose output is fused afterwards with the other modalities. The module can then utilise all common text processing methods, such as stemming and part-of-speech tagging, and any of the existing tools for text-based indexing. In the experiments of this paper, we use textual search for indexing text obtained with automatic speech recognition and machine translation.

2.4 Concept detection

After having extracted low-level video features from each shot, supervised learning techniques are applied for learning the associations between the low-level features and the concepts in the annotations of the video corpus. Figure 2 illustrates the overall architecture of the concept detection module of the PicSOM video database search system. In this architecture, concepts are first detected independently from each video shot. Shot-wise concept detection is achieved by fusing together the outcomes of numerous elementary detectors, each of which is based a different visual feature of the shot. The feature-wise elementary detectors build upon the support vector machine (SVM) supervised learning algorithm. See [3] for more details.

Following the shot-wise detection phase, the detector outcomes are post-processed so that advantage is taken of the correlations that contents of temporally adjacent video shots often exhibit. Technically, this is achieved via concept-wise inter-shot N-gram modelling [6].

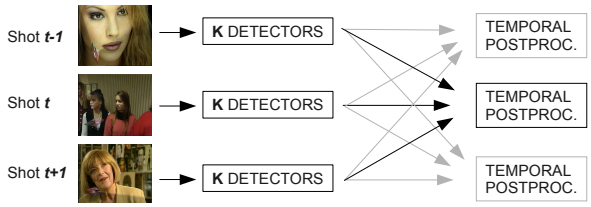


Figure 2: Overview of concept detection.

2.5 Video search

The goal of video retrieval is to find relevant video content for a specific information need of the user. The conventional approach has been to rely on textual descriptions, keywords, and other meta-data to achieve this functionality, but this requires manual annotation and does not usually scale well to large and dynamic video collections.

Content-based video retrieval, on the other hand, utilises techniques from related research fields, such as image and audio processing, computer vision, and machine learning, to automatically index the video material with low-level features. Content-based queries are typically based on provided examples (i.e. *query-by-example*). The video collection is ranked based on its similarity to the examples. See [2] for a description of the content-based retrieval methodology employed in the PicSOM system.

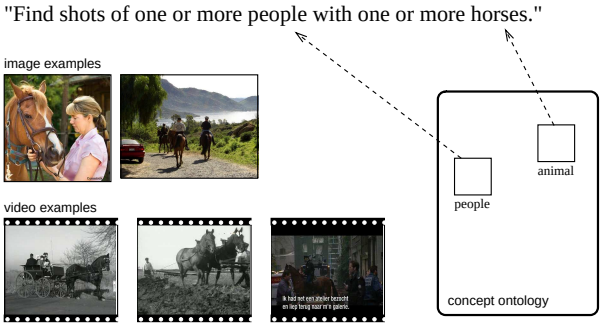


Figure 3: An example TRECVID search topic, with a possible concept mapping from a concept ontology.

In concept-based video retrieval, the fundamental problem is how to map the user's information need into the space of available concepts in the used concept ontology. The basic approach is to select a small number of concept detectors as active and weight them based either on the performance of the detectors or their estimated suitability for the current query. We have employed two methods for automatic concept selection: *text-based* and *visual-example-based*. The text-based approach uses lexical analysis and synonym lists to match the words of the textual query to the available concepts. In the visual-example-based method the PicSOM system measures the similarity of the given visual examples (images and/or videos) to the concept detectors, answering the question of how visually similar are the examples to the annotated training set. See [3] for details.

3. EXPERIMENTS IN TRECVID

The video material and the search topics used in these experiments are from the TRECVID evaluations [5] in 2008–2009. TRECVID is an annual workshop series organised by the National Institute of Standards and Technology (NIST) and arguably the leading venue for evaluating research on content-based video analysis and retrieval. It provides the participating organisations large test collections, uniform scoring procedures, and a forum for comparing the results. Each year the TRECVID evaluation contains a set of video analysis tasks, such as semantic concept detection, video search, video summarisation, and content-based copy detection. In the experiments of this paper, we focus on the settings of the automatic video search task.

In 2008–2009 the type of video material used in TRECVID consisted of documentaries, news reports, and educational programming from Dutch TV. The video data is always divided into separate development and test sets. The same development set of approximately 100 hours in length is used both in 2008 and 2009. The amount of test data was approximately 100 and 280 hours in 2008 and 2009, respectively. The training data for semantic concept detection is obtained with a collaborative annotation effort among the participants.

NIST also defines sets of standard search topics for the

video search tasks and then evaluates the results submitted by the participants. The search topics contain a textual description along with a small number of both image and video examples of an information need. Figure 3 shows an example of a search topic, including a possible mapping of concept detectors from a concept ontology based on the textual description. The number of topics evaluated for automatic search was 48 in the year 2008 and 24 in 2009. Due to the limited space, the search topics are not listed here, but are available in the TRECVID guidelines documents[1].

The video material used in the search tasks is divided into shots in advance and these reference shots are used as the unit of retrieval. The shot segmentation step resulted in 36 000 shots for the training set and 97 000 shots for the full 2009 test set. The output from an automatic speech recognition (ASR) software is provided to all participants. In addition, the ASR result from all non-English material is translated into English by using automatic machine translation. It is therefore quite unsurprising that the quality of the textual data is remarkably poor and pure text queries can only obtain a very modest performance.

Search performance in TRECVID is measured using *mean average precision* (MAP) and *mean inferred average precision* (MIAP). MIAP is an approximation of MAP, but requires only a subset of the results to be evaluated manually.

In Figure 4 the automatic search performance of the current PicSOM system is compared with the top automatic search systems submitted to TRECVID 2008 and 2009. Our own submissions in those years are included as well (with the year indicated). We see that the PicSOM system compares very well with the other top systems.

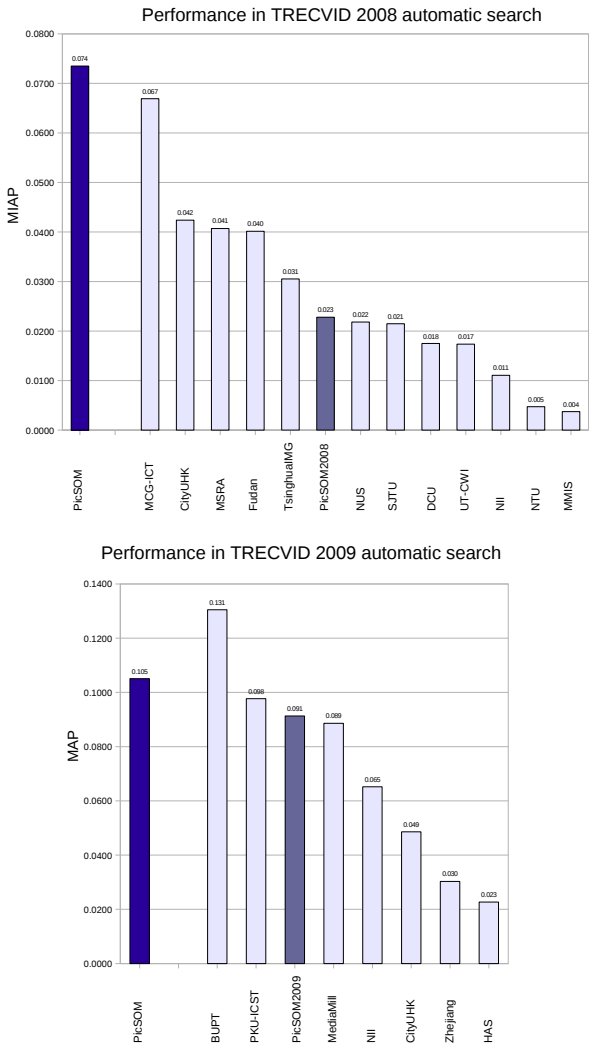


Figure 4: MIAP/MAP performance in TRECVID

4. CONCLUSIONS

This paper has given an overview of video search using the PicSOM system utilising concept-detection for helping to bridge the semantic gap between the low-level features and the high semantic level of a human query. By comparing with other state-of-the-art systems in the TRECVID 2008 and 2009 competitions we have demonstrated the competitiveness of the proposed retrieval method for automatic search.

Automatic video search systems can find practical application in web-wide search engines and in the public archives of a TV station or other content-producing institutions. For example, the data set of previous TRECVID evaluations have been TV broadcasts and the upcoming TRECVID 2010 data comes from the Internet Archive's collection of public domain videos. Providing efficient access to the growing amount of visual data created today is not only a matter of convenience, but also crucial for the preservation of our contemporary culture. Only information that can be found and retrieved with reasonable effort is useful in practice.

5. REFERENCES

[1] A. G. Hauptmann, M. G. Christel, and R. Yan. Video retrieval based on semantic concepts. *Proceedings of the IEEE*, 96(4):602–622, April 2008.

[2] J. Laaksonen, M. Koskela, and E. Oja. PicSOM—Self-organizing image retrieval with MPEG-7 content descriptions. *IEEE Transactions on Neural Networks*, 13(4):841–853, July 2002.

[3] M. Sjöberg, V. Viitaniemi, M. Koskela, and J. Laaksonen. PicSOM experiments in TRECVID 2009. In *Proceedings of the TRECVID 2009 Workshop*, Gaithersburg, MD, USA, November 2009.

[4] A. F. Smeaton. Techniques used and open challenges to the analysis, indexing and retrieval of digital video. *Information Systems*, 32(4):545–559, 2007.

[5] A. F. Smeaton, P. Over, and W. Kraaij. Evaluation campaigns and TRECVid. In *Proc. of the 8th ACM International Workshop on Multimedia Information Retrieval*, pages 321–330, New York, NY, USA, 2006.

[6] V. Viitaniemi, M. Sjöberg, M. Koskela, and J. Laaksonen. Exploiting temporal and inter-concept co-occurrence structure to detect high-level features in broadcast videos. In *Proceedings of WIAMIS 2008*, pages 12–15, Klagenfurt, Austria, May 2008.

[1] *http://www-nlpir.nist.gov/projects/trecvid/*

ACKNOWLEDGEMENT

This work has been supported by the Aalto University MIDE programme, project UI-ART.

Personalisation Technologies for IPTV channels. Mashup-TV project.

Toni Bibiloni, Gabriel Fontanet, Jorge Chao, Cristina Pérez
Lab. Multimedia - University of Balearic Islands
Ctra. de Valldemossa, km 7.5
Palma de Mallorca, 07122
+34 971173204
{toni.bibiloni, gfontanet, jorge.chao, cristina.perez}@uib.es

ABSTRACT
Mashup-TV project is a system that combines audiovisual content and data information from multiple external sources in order to create a new service, called "Personalise IPTV touristic channel". That project implies easily a fast integration and benefits from data sources that never were created for the tv-channel production. For example, data sources from administration, government, etc., data information from enterprise and intelligent use of ads.

Mashup-tv lies in the richness of interactive functionality as well as the possibility for personalised services and the chance to distribute own media among the "TV tourist community".

"Tourist" consumer acceptance of IP-based television dramatically changes the distribution of audiovisual content in private networks, in our case, hotels. New models for media, software and services are emerging, and this involve the audience itself. This paper focus on aspects of web services, semantic web, enrich IPTV , SMIL 3.0, MPEG-7 and full interaction and personalisation.

Categories and Subject Descriptors
H.3.1, H.3.4, H.3.5 [**Information Storage and Retrieval**]: Content Analysis and Indexing – *Indexing methods;* Systems and Software – *User profiles and alert services;* Online Information Services – *Web-based services;* H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Input devices and strategies;*I.2.4 [**Artificial Intelligence**]: Knowledge Representation Formalisms and Methods – *Ontologies.*

General Terms
Algiorithms, Management, Design, Human Factors.

Keywords
User profiles, information management, personalization, Ontology, Semantic Web, metadata, multimedia index structures, MPEG-7, multimedia presentation, mashup, SMIL, interactive DTV, tourism management.

1. INTRODUCTION
A broadband mashup-TV is the combination of multiple sources of video, images and text information which initially have no relation between each other, but the are actually linked by location, thematics, user and hotel profiles, etc.

This paper illustrate how Semantic Technologies can support information integration and make it easy to create Personalised Mashup-TV channels. We present an Mashup-TV touristic project based on an intelligent ontology, metadata MPEG-7 and interactivity over SMIL presentations.

The Mashup-tv project offer to consumers a personalised tv production with interactive and hyperlinked functionalities over the rich audiovisual media transmitted over IP protocol. The mashup-tv touristic channel offer attractive content for specific interest tourists groups, always combined with added information of advertising, timetables, price list and the possibility of booking. Interactive booking adds an attractive extra value to the user experience and it is widely accepted by the audience.

Back-office mashup-tv offer building online productions through mashing audiovisual content in real time from different providers and servers, using web services to add dynamic information, using open APIs and data sources to produce results with semantic relations. For example, used of cartographic data to add location information to promotional videos. Mashup-tv combine from multiple sources (videos, timetables, ...) into a single representation using SMIL 3.0 for offer interactivity and final composition over the television set. The used content is typically sourced from a third party via a public interface.

For content provider, in our case, the hotel, new possibilities to sell "services" to consumers involve additional benefits. Personalised audiovisual units offer huge opportunities for private IPTV networks. Business Mashups are useful for integrating business and data services, as Business Mashups technologies provide the ability to develop new integrated services quickly, to combine internal services with external or personalised information, and to make these services tangible to the business user through user-friendly browser interfaces [5].

The mashup-tv project offer the chance to create a completely new service for each customer. This is achieved by a conceptual user's profile model. The hotel can access to the personal customer's information, as preferences, needs and interests that are recorded in a user profile ontology and assigned to a room number.

We begin in sections 2 with description of technologies. Section 3 present the details and architecture. In section 4 case study: "hotel scenario". Section 5 conclusions.

2. TECHNOLOGY MOTIVATION
Mashup-TV project means content aggregation technologies, personalised audiovisual production plus intelligent interactivity

for useful client audiovisual presentation. There are several challenges:

1. Integration of heterogeneous online travel information. Develop APIs to provide content from external sites, mashup technology.
2. Semantic matching tourist services and user preferences. Operate on pure MPEG-7 content and ontologies
3. Profiles.
4. Create audiovisual presentation-oriented content. Involve SMIL 3.0.

2.1 Integration

It involves a large amount of detailed up-to-date information. The information is available over the Internet, however, due to the heterogeneity of the information, it is difficult to automatically integrate the information.

The aim of the Mashup-tv is to give the audience a flexible and targeted access to a particular interest and services. The user can navigate and access to the topics that are geographically close to the hotel.

2.2 Semantic Database

One way to model a domain ontology is to reuse o extend an existing ontology, we expanded the Ontour Ontology [7] and we create the Mediatur Ontology [8]

One crucial point is the use of metadata for a targeted access to semantic relations between particular audiovisual units and "hyperlinks" in Internet, in order to offer a intelligent audiovisual presentation and interaction with content. [4] The metadata database is MPEG-7 compliant. Metadata play a very important role for the navigation, particularly for the subject, user preferences and geographical location [6].

The project is able to answers various kinds of questions that can be asked when the user interact with Mashup-TV channel. These questions involve the predicates What, Where, When, and How.

2.3 User Profiling

Currently, all users use services on the internet or even in desktop applications where they must register to access personalised services. These services represent a greater or lesser extent, computational models of users, a number of stored data corresponding to a virtual person is model in the system and that identifies them.

As shown in [2], user information to build profiles or models can be obtained directly through the data entered by the user in the system and the data obtained indirectly through relationships with other systems, in order to take a complete representation of the user and to offer services that fit your preferences.

To complete these user profiles and categorize them, it needs a wide variety of personal data and preferences, as shown in [3]. In the case discussed in this paper, we must have two different types of profiles:

- Hotel profiles
- Tourist profiles

Hotel profile should have special features for the hotels classification and customize the content displayed to your customers. For example, if a hotel is located in beachfront, your profile will not be in rural areas, but will be allocated in a specific tourism model.

Moreover, tourist profile will include of a wide variety of personal data, through which you can categorize and deliver content completely customized and tailored to their personal preferences or personal context during the time of staying at the hotel. For example, if the tourist is staying at the hotel to which we referred previously and he likes surfing, the system must be able to offer a place to buy a new surfboard or where to buy surf clothes near the hotel location.

User profiling is very important to offer a personal pleasant services and offer a high quality service in a tourism environment.

2.4 Smil 3.0

One of the topics of this project is use SMIL 3.0 in digital TV services. Nowadays, people can not think in web environment without multimedia concept. Anytime and everywhere, people needs search and share all the information that they require: text, images, videos, audios. Join all this contents in a presentation where the user can have them all, can be very complicated. SMIL makes it easier.

In December 2008, W3C published the new SMIL 3.0 draft, updating the previous version 2.1. But, what is new on it? Referring to [1], the main improvements that we highlight are:

- SMIL 3.0 smilText provides a new media type for use in SMIL presentations. Unlike other media types defined in the media object module, smilText provides a text container element with an explicit content model for defining in-line text, and a set of additional elements and attributes to control explicit in-line text rendering. This module is very important for the subject of this paper, where accessibility and text rendering are a critical point in a multimedia presentation.
- SMIL 3.0 State provides a mechanism for the author to create more complex control flow than what SMIL provides through the timing and content control modules, without using a scripting language. To provide this, it allows a document to have some explicit state (variables) along with ways to modify, use and save this state. This new module enables greater interactivity and flexibility of the multimedia applications. We will use to improve interactivity in audiovisual productions.

In composition, the most important are:

a) The BasicLayout module is extended with the backgroundOpacity attribute, which specifies the background opacity of a region. This attribute is very important for the layout design, because it allows regions that won't be shown completely but defined to add multimedia content and visualize it.

b) New MediaPanZoom module for panning and zooming over media content and MediaOpacity module containing new rendering attributes for chroma key and opacity control. These modules allow more design functionalities, like camera motion or Ken Burn effects.

c) One of the most important problems has SMIL was text rendering. In the last version, it was created a new BasicText module provides basic support for in-line text within a SMIL presentation; a new TextStyling module that adds styling capabilities of text and TextMotion module defines how text is added. Improve the capacity of design makes it get better presentations and present information more understandable.

3. ARCHITECTURE

The proposed system architecture is shown in figure 1. It's divided into four parts. Firstly, it is the metadata created using MPEG-7 descriptions of audiovisual content (indexing proces) and apply it over a toruist ontology. Secondly, the information repository of ontology is used to create a tourist-TV channel (contextual proces). Thirdly, the server hotel execute the APIs mashup and preference profile to offer personalized tourist channel. Finally, the customer interact over the STB.

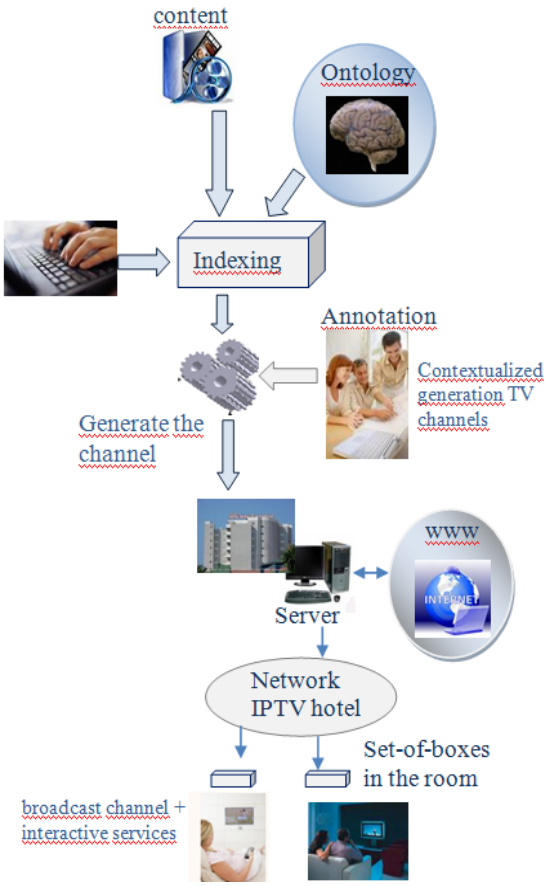


Figure 1. MASHUP-TV System Architecture

4. HOTEL SCENARIO

The first analyses of behaviour on use IPTV systems shows a strong demand of VOD services, but in the "tourist IPTV channel" the user will be able to decide between topics, full control over the selection, intensive use of interactive services, participate, etc.

The aim of the project is to give to the "tourist" a flexible and as much as possible targeted access to corresponding media, even shopping advises as well as audiovisual advertising units. The system knows tourist preferences, and then show content and additional information, it can changes permanently over time o desired location. The project behaves like a recommendation system.

The system offer answers to typical questions, some examples:

- What can a tourist do while staying tomorrow morning in the hotel.
- Where are located the interesting places to see and visit to go walking from the hotel.
- When can the tourist visit a particular place, the price, public transport?
- How can the tourist get to its destination to do an activity, like wind-surf?

An example is shown in figure 2, the audiovisual presentation over IPTV terminal can combine audiovisual media from the repository, integration of many other web services, Google Maps for appropriate maps, book a holiday service, to book a restaurant, etc.


```
<smil xmlns="http://www.w3.org/ns/SMIL" version="3.0"…>
<head>
<textStyling>
…
</textStyling>
<layout>
<root-layout xml:id="principal" …" />
<region xml:id="multimedia" …/>
…
</layout>
…
</head>
<body>
…
<par>
<smilText region="titulo" textPlace="center">
Monument - Palma's Cathedral
</smilText>
<img dur="13"region="multimedia" transIn="Transicion2"…>
<area href="http://www.catedraldemallorca.org"
sourcePlaystate="play" external="true" acceskey="1"/>
</img>
```

```
<img src="http://maps.google.com/staticmap..." />
<smilText region="subtitulos" …>
Facade of the cathedral of Palma de Mallorca
</smilText>
<audio src="…" dur="00:00:04" region="subtitulos" />
</par>
…
</body>
</smil>
```

Figure 2. Example of Audiovisual presentation over IPTV and SMIL 3.0 codegenerated

5. CONCLUSIONS

In this paper we proposed an ontology and an intelligent information recommendation system using a IPTV channel. The system offers the personalized recommendations of tourist attractions. Through the ontology of tourism, the system allows integration of heterogeneous online travel information.

Mashup-TV is based on Web 2.0 Technologies, advanced features built on HTML, Javascript and CSS technologies. Optimum flexibility and control for using RTSP protocol. Enables cost-effective design using existing web-based skills. Content providers are integrated using APIs. middleware companies. Multi-lingual support built in, with new languages easily added.

6. ACKNOWLEDGMENTS

This research was partially supported by TSI-020110-2009-26 project under grant of MINISTERIO DE INDUSTRIA, TURISMO Y COMERCIO, Gobierno Español, Spain [9].

7. REFERENCES

[1] W3C Recommendation SMIL 3.0. Available: http://www.w3.org/TR/2008/REC-SMIL3-20081201/

[2] Ghost, R. and Dekhil, M.2009.Discovering User Profiles. Poster Sessions. Hewlett-Packard Laboratories. (Madrid, Spain, April 20 - 24, 2009). DOI= http://doi.acm.org/10.1145/1526709.1526944

[3] Ghost, R. and Dekhil, M.2008.Mashups for Semantic User Profiles. Poster Sessions. Hewlett-Packard Laboratories. (Madrid, Spain, April 20 - 24, 2009). DOI=http://doi.acm.org/10.1145/1367497.1367740

[4] Lugmayr, A. Niiranen, S. Kalli, S. 2004. Digital Interactive TV and Metadata - Future Broadcast Multimedia. Berlin. Springer.

[5] Mashup business scenarios and patterns: Part 1. Avaible: http://www.ibm.com/developerworks/lotus/library/mashups-patterns-pt1/

[6] Wang, W, Zeng, G., Zhang, D., Huang, Y., Qiu, Y., Wang, X. 2008. An Intelligent Ontology and Bayesian Network based Semantic Mashup for Tourism (IEEE Congress on Services 2008).

[7] Prantner, K., Ding, Y., Luger, L., Yan, Z. 2007.Tourism ontology and semantic management system: State-of-the-arts analysis (IADIS International Conference WWW/Internet 2007).http://www.iadis.net/dl/final_uploads/200712C070.pdf

[8] Mediatur Project. New Experiences in tourism management content. Available: http://www.mediatur.net/

[9] Mashup-TV project. Available: http://caval.travel/projects/mashup-tv/objetivos.html

Media Content Management Platform

Francesc Pinyol Margalef
GTM – Grup de Recerca en Tecnologies Mèdia, LA SALLE – UNIVERSITAT RAMON LLULL
Quatre Camins 2, 08022 Barcelona, Catalonia, Spain
fpinyol@salle.url.edu

Marc Rovira Vall
GTM – Grup de Recerca en Tecnologies Mèdia, LA SALLE – UNIVERSITAT RAMON LLULL
Quatre Camins 2, 08022 Barcelona, Catalonia, Spain
marcr@salle.url.edu

Jordi Borrell
GTM – Grup de Recerca en Tecnologies Mèdia, LA SALLE – UNIVERSITAT RAMON LLULL
Quatre Camins 2, 08022 Barcelona, Catalonia, Spain
jordibs@salle.url.edu

Gabriel Fernandez Ubiergo
GTM – Grup de Recerca en Tecnologies Mèdia, LA SALLE – UNIVERSITAT RAMON LLULL
Quatre Camins 2, 08022 Barcelona, Catalonia, Spain
gabrielf@salle.url.edu

ABSTRACT
The Media Content Management Platform (MCMP) is a wide standard-based media content cataloguing and managing system which objective is the uploading, searching and viewing of on-demand videos and TV channels. Currently, some cataloguing systems do not store enough information about their contents, and this content data does not follow any metadata standard, so they do not use an interoperable metadata model. On the contrary, MCMP stores in its database a large list of media content metadata based on standards such as MPEG-7, MPEG-21 and TV-Anytime. It allows having an extensive and complete media content description with a standard and interoperable system. This metadata is structured in core metadata, semantic metadata and technical metadata. Thanks to this approach content customization according to user preferences, better searching system and an extensive and comprehensive content information for the user can be provided. All this capabilities raise the complexity of media content management, but it does not decrease the usability, otherwise it provides to users and administrators more tools and possibilities to manage media content data more efficiently and effectively. The system allows user-generated or professional videos, to be streamed with RTSP, or television channels described by SDP files.

Categories and Subject Descriptors
H.5.1. [**Information Interfaces and Presentation**]: Multimedia Information Systems – *video*.

General Terms
Management, Human Factors, Standardization.

Keywords
MPEG-7, MPEG-21, TV-Anytime, video, media, catalogue, streaming, metadata.

1. INTRODUCTION
The exponential increment of media contents in the World Wide Web causes the need to have better cataloguing systems for these contents in order to improve content searching and user customization.

Nowadays two paradigms about exchange of media contents in internet can be found: the web paradigm and P2P paradigm. The main representatives of web paradigm are YouTube and Google Videos. These platforms offer the possibility to upload, search and view all kinds of videos. These platforms are oriented to non expert users who need an easy way of manage videos, using in this case a few properties to catalogue media content as title, description, tags and category.

The other paradigm, P2P, has as main representatives Emule and BitTorrent. They offer the possibility to download, search all kinds of videos and other media content. These platforms are oriented to share all kind of files, and they are not focussed only in media content. Furthermore, the associated content metadata is insignificant: only the title, which is insufficient for advanced media content management.

In order to solve this situation, a complete and standard media content management system has been designed. It is complete because it joins up all the semantic and technical metadata needed to achieve a full data model for an advanced media content management. It is standard because it follows MPEG-7 [1] and MPEG-21[2] metadata standards as data model.

These standards offer the chance of choice a wide range of media properties (MPEG-7) as semantic and technical data, and a way to encapsulate this metadata in a coherent XML tag structure. Finally, this standard solution lets us to be compatible and interoperable with other similar media content management platforms in order to compare, share and distribute media content.

This media content manager is built as a web service. This characteristic is essential to achieve an easily combinable module with other modules or managers present in a whole media system.

So, content manager is a part of distributed system, which can be called from other distributed managers, as user manager, to conceive a larger software structure.

Finally, all media content added in media content manager repository is transcoded to all media content manager platforms. These platforms are PC, IPTV, DTV and Mobile. Each platform has its own media format, adapted to the user device features.

The rest of this paper is organised as follows. Section 2 provides an overview of whole project context. Section 3 provides a deep introduction of media content manager module. The details of selected data model are elaborated in Section 4. Section 5 gives a detailed description of web service. Section 6 provides an overview of web application, followed by the concluding marks in Section 7.

2. PROJECT CONTEXT

The MCMP is a part of a parent project called RAUDOS [3]. It is a research and development project, which goal is the design and development of a multiplatform interactive system for media content distribution. The novelty of RAUDOS is that the user by interacting with the system can earn points, which can be exchanged for media contents.

The business model of parent project consists of getting advertising revenue managed by an advertising manager. The advertising manager, together with the customization manager, may select the advertising profiled to the user preferences.

The project core consists mainly of a number of repositories and managers (see Figure 1), which are working distributed and connected by web services.

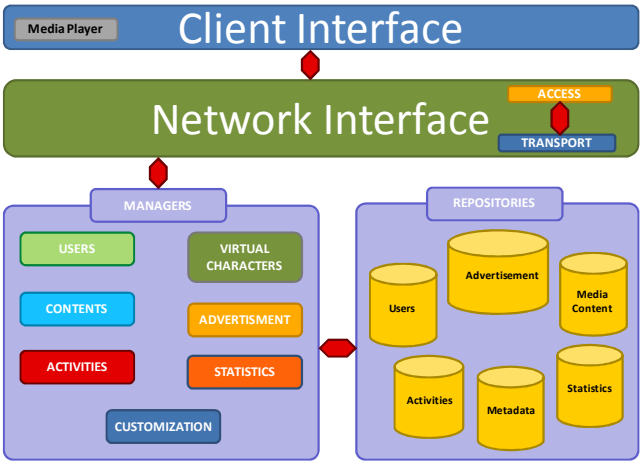


Figure 1: Project architecture.

- **User Manager**: this manager is responsible for the authentication of users on the platform. It also identifies the platform from which they access.
- **Content Manager**: this manager is responsible for providing all audio-visual resources. These resources are fully indexed so as to facilitate searches.
- **Activity Manager**: this manager, based on the history of actions, consumer and user profile, can select content and advertising tailored to their preferences.
- **Customization Manager:** according to the history of actions, consumer and user profile, customization manager can search, sort and present the most appropriate content and advertisement to users, based on their tastes and preferences.
- **Publicity Manager**: this manager can add, edit or delete banners of advertisements, both static and dynamic, and allows credit increment when they are consumed.
- **Virtual Characters Manager**: this manager handles the appearance and behaviour of virtual characters that interact with the user so that the user interaction with system becomes smooth and pleasant.
- **Statistics Manager**: this manager is responsible for storing the key performance indicators. It enables the creation of a scorecard for system management.

This article describes in detail the content manager.

3. CONTENT MANAGER MODULE

The content manager is responsible for storing content and metadata in the respective repositories or databases. Afterwards, these media contents can be edited, deleted and/or displayed.

More precisely, users of content manager can perform some specific tasks:

- **Add Content**: users can upload content to the content repository. This content can be a video or a television channel. If it is a video, it is transcoded to all platforms formats. In both cases, users can add the content metadata: core, semantic and technical properties.
- **Distribute Content**: it allows the dissemination of content for their consumption and visualization using RTSP [4] for streaming video.
- **Distribute Metadata**: users can get the technical, semantic and core metadata associated with content.
- **Search Contents**: users receive a list of content from the search and sort criteria of metadata.
- **Edit Content**: users can modify the core, technical or/and semantic metadata of selected content.
- **Delete Content**: users can delete content and its metadata from platform.

The actors who perform the tasks are:

- **Customization Manager**: it is used by the content manager to search for contents and distribute their metadata.
- **Consumer**: the responsible for consuming the contents. it needs to search the contents. Then it can see the metadata associated with content. Finally, he or she can request the distribution of the content, to be consumed in streaming.
- **Administrator**: the responsible for overall management of system platform. Therefore, he or she may take all actions

such as add, edit, delete videos and TV channels, search, and distribute their metadata and content.

Content manager contains a media content repository, where users can upload their videos and TV channels files, a metadata repository with a database with all media content properties.

4. DATA MODEL

Media contents have associated a large number of properties or metadata. These data can be semantic information about the content as title, creators, genre, abstract...; information about media file as duration, size, codec...; and core information as "raudollars" (tokens), recommendation value, content type and thumbnail.

These metadata are defined and organised according to MPEG-7 and MPEG-21. MPEG-21 offers a standard system of organization and encapsulation of the metadata content on Digital Items. MPEG-7 gives us the definition of metadata, both technical (see Table 1), and semantic (see Table 2).

For dealing with genre and subject, which categorises the content that the user or administrator brought into the manager, the TV-Anytime standard has been relied on [5], from which the subset that is more representative for cataloguing the contents has been taken.

The content search is performed through a search and sort criteria of results. Search criteria are based on the definition of metadata, adding the ability to create ranges or limits. The sorting allows us to quantify results and sort them by fields.

The contents are identified in the database using a Universally Unique Identifier [6] (UUID). It is a good system for distributed communication between several managers of system.

Table 1: Technical metadata.

MPEG-7
FileFormat
MediaDuration
Visual coding - Format
BitRate
Visual coding - Pixel - BitsPer
Visual coding - Frame width
Visual coding - Frame height
Visual coding - Pixel aspectRatio
Visual coding - Frame rate
FileSize
Audio coding.-Format
Audio coding - Sample - Rate
Audio coding - Sample - BitsPerPixel
Audio coding - AudioChannels

Table 2: Semantic metadata.

MPEG-7
Title
Genre
Subject
Language
CaptionLanguage
AttributeValuePair
AttributeValuePair + Temporal Descomposition
CreationCoordinates Date
CreationTime
ReleaseDate
Keywords
Abstract
Target Market
Media Review
Creator + Rol
Who
WhatAction
Where
When
WhatObject
Why
How

5. WEB SERVICES

Web interfaces and all managers of system need to communicate among them in a distributed way. With this intention content manager has been built as a web service, which communicates via SOAP with others managers.

Java has been used as a programming language and JBossWS-3.0 [7] as a web service library. This application works in a JBoss Server 4.2.3 [6]. MCMP uses Hibernate tool [8] in order to map relational database fields in a class model of application.

Content Manager Web Service (CMWS) has the same functionalities as tasks of content manager module (see Table 3).

AddContent saves video file of content in *Incoming* folder of content repository. AddContent also saves initial metadata of content in database. The content in *Incoming* folder is transcoded to all platforms: PC, IPTV, DTV and Mobile.

Table 3: Web service functions.

1	String **getContent** (int platform, String contentId)
2	String **addContent** (String location)
3	String **getTVChannel** (String contentId)
4	String **addTVChannel**(String sdpLocation)
5	void **editContent**(String contentId, MetadataInfo metadata)
6	MetadataInfo **getMetadata** (String contentId, int metadataSet)
7	void **deleteContent** (String contentId)
8	List< String> **search** (SearchCriteria searchCriteria, SortCriteria sortCriteria)

AddTVChannel saves an SDP [9] file of TV channel in *sdps* folder, and it also saves initial metadata in the database.

GetMetadata and editContent work with MetadataInfo data structure. This data structure contains all technical metadata, semantic metadata and core metadata of system. We also can

appreciate metadataSet integer, which is useful for choosing a subset of requested metadata.

The Content Manager search allows basic search and advanced search. Basic Search looks in the main fields of the database from a given text, whereas Advanced Search looks in all specified fields of database.

6. WEB APPLICATION

MCMP has a web user interface, which is used to perform the tasks of the CMWS via Web.

This software layer is a PHP-HTML-JAVASCRIPT [10] application, with the following web interfaces: contentUpLoad, metadataView, editMetadata, contentSearch, contentList catalogue, contentDetail , contentWatch.

Figure 2: Catalogue interface.

These interfaces communicate with different managers via SOAP and are able to perform all content manager tasks. Users can upload media content to the server, and then access to a page form in order to add the metadata.

Once metadata are submitted, users can access to them: they can perform content searches.

The user can delete the selected content, buy it, or view all its metadata. When a user buy the content, his or her credit is reduced and the user is redirected to a page in order to consume it.

In Figure 2 we show a screenshot of the Catalogue graphical interface, where the user can introduce semantic as well as technical metadata related with the uploaded content.

7. CONCLUSIONS

MCMP is a complete media content management system for professional or advanced environments, which allows us having a standard interoperable system, thanks to MPEG-7, MPEG-21 and TV-Anytime standards.

Furthermore, this system works as a multiplatform and multidevice system, because it can transcode the videos to different platforms or devices.

8. REFERENCES

[1] P. van Beek, A. B. Benitez, J. Heuer, J. Martinez, P. Salembier, Y. Shibata, J. R. Smith, 2001. Information technology — Multimedia Content Description Interface — Part 5: Multimedia Description Schemes. ISO/IEC JTC 1/SC 29/WG 11/N3966.

[2] R. Van de Walle, I. Burnett, G. Drury, 2003. Information Technology – Multimedia Framework (MPEG-21) – Part 2: Digital Item Declaration (Second edition). ISO/IEC JTC 1/SC 29/WG 11/N6164.

[3] E. Burgoa, M. Rovira et al., 2010. RAUDOS: Interactive Multiplatform Audiovisual Delivery. In Proceedings of Euro ITV 2010, in press.

[4] H. Schulzrinne, A. Rao, R. Lanphier, 1998. Real Time Streaming Protocol (RTSP). http://tools.ietf.org/html/rfc2326

[5] TV-Anytime Forum, 2001. Specification Series: S-3: Metadata. Appendix B TV-Anytime Genre Dictionary. http://xml.coverpages.org/ni2001-10-02-c.html

[6] Open Source Software project (OSSP). Universally Unique Identifier (UUID) http://www.ossp.org/pkg/lib/uuid/

[7] JBoss Community, 2007. JBoss Server Manager Reference Guide Version: 2.0.0.GA. http://docs.jboss.org/tools/2.0.0.GA/as/en/pdf/server_manager_guide.pdf

[8] G. King, C. Bauer, M. Rydahl Andersen, E.Bernard, and S. Ebersole, 2009. Hibernate Reference Documentation 3.3.2.GA. JBoss Community. http://docs.jboss.org/hibernate/stable/core/reference/en/pdf/hibernate_reference.pdf

[9] M. Handley, V. Jacobson, 1998. SDP: Session Description Protocol. http://tools.ietf.org/html/rfc2327.

[10] Arnaud Le Hors, Ian Jacobs, 1999. HTML 4.01 Specification. http://www.w3.org/TR/1999/REC-html401-19991224/

RAUDOS: Interactive Multiplatform Audiovisual Delivery

Eduardo Burgoa
AIDO
Nicolás Copérnico, 7-13, Paterna
46980 VALENCIA SPAIN
+34 961 318 051
eburgoa@aido.es

Marc Rovira
LA SALLE
Quatre Camins 2, 08022
BARCELONA SPAIN
+34 932 90 24 28
marcr@salle.url.edu

Fernando García
ANDAGO
Alcalde Ángel Arroyo, 10, 1º planta
Getafe, 28904 Madrid, SPAIN
+34 916 01 13 73
fernando.garcía@andago.es

ABSTRACT
This paper proposes a new multi-device and multi-technology audiovisual content delivery system. Its main features are interactive applications, content and advertisement recommendation, and virtual character guidance.

Categories and Subject Descriptors
H.5.1 [**Multimedia Information Systems**]: Video distribution system

General Terms
Experimentation

Keywords
Audiovisual, Interactivity, Multi-platform, DVB-T, DVB-H, IPTV, Kiosk, Virtual Characters, MPEG-7, MPEG-21, MHP

1. INTRODUCTION
RAUDOS is a novel multiplatform interactive audiovisual content delivery system that is accessible and adapted to every user at a competitive cost for businesses and consumers.

Any RAUDOS user connects to the Platform with its user name and password. Once inside, user interacts with entertainment and educational applications that allows it to earn virtual money with the virtual character advice. This kind of interaction can be done with any sort of device connected to the network (computer, mobile phone, digital TV Set-Top Box or virtual kiosk). User has access to an audiovisual content catalogued in order to buy content with its virtual money (called *tokens* or *raudollars*). This audiovisual content can be seen in any of the above mentioned platforms.

The system encourages the authorized audiovisual content consume without real money cost for the final client using novel advertisement business models.

The rest of this paper is organized as follows: section 2 describes the overall system architecture; section 3 gives an overview of the main access platforms; sections 4, 5, 6, 7, 8, 9 explain main innovations; and future work is given in section 10.

2. ARCHITECTURE
RAUDOS system can be divided in two main blocks: *The Core system* and *The Access platforms*.

The Core system is formed by a logic tier and a data tier. Access platforms define presentation tier as well as the user interface to the system.

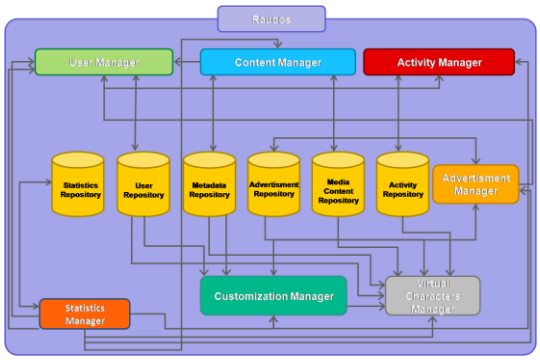


Figure 1: RAUDOS Architecture

User Manager: This module is in charge of user authentication in the Platform as well as the maintenance of user session and identification throughout the Platform.

Content Manager: This module manages the RAUDOS audiovisual content resources.

Activity Manager. This module is in charge of interactive activities offered by RAUDOS.

Customization Manager. The audiovisual content and the advertisement recommendation are administered by Customization Manager.

Advertisement Manager. This module is in charge of advertisement information of the system.

Virtual Character Manager. Appearance and behaviour of virtual character that interact with user is managed by this module.

Statistics Manager. This module is in charge of maintaining the key performance indicators of RAUDOS.

Communications between logic tier and presentation tier are performed by the service interface, implemented with HTTP/SOAP web services, offered for each module of the system.

On the other hand, RAUDOS offers access to services among a lot of different technologies: DVB-T, DVB-H, IPTV, 3G/UMTS/HSPA, and virtual kiosks.

3. PLATFORMS

3.1 DVB-T Platform

The small chance of having a return channel for DVB-T using the same platform of the broadcast media, makes the community to find a solution (though there are solutions such as DVB-RCT standard, they have no market or infrastructure). The idea that has been adopted in the RAUDOS architecture is the use of the capabilities offered by the MHP standard over interactive applications on digital television. Using the broadcast channel, an interactive application is delivered together with conventional audio and video streams of digital television (also known as Xlet, Java-based and executed by set-top-boxes). This application is a XHTML browser based on X-Smiles [1] and is received and executed by the STB. It allows the user to display on its TV conventional web pages that follows the XHTML standard, getting dynamic pages in order to access web services that are published by various managers of the RAUDOS architecture. These requested pages are sent through the Internet channel and they are customized to the profile of the user and the STB features.

3.2 Mobile Platform

The Mobile Platform is responsible for adapting the services and functionalities offered by the RAUDOS core to the peculiarities of mobile devices. These peculiarities are mainly the small screen size, the relative difficulty of data entry, low processing speed, RAM quite small and a limited internet bandwidth.

To achieve this goal, the platform provides a simplified web interface to be used from the browsers that current mobile devices incorporate (phones, PDAs, etc.). The browser will use an external video player for viewing the RTSP and RTP video streams offered by the RAUDOS core. For the transport of HTTP/IP (web interface) and RTSP/IP (unicast video streams) data, which by its nature is bidirectional, the mobile device makes use of 3G/UMTS/HSDPA technology, and for the transport of the RTP/IP broadcast data (TV channels) makes use of DVB-H technology, which allows to offer video services to large numbers of people on a scalable manner and with reduced cost.

The interface of the mobile platform consists of a series of screens that provide access to RAUDOS core services: new user account, access (login), menu, profile, activities, videos (RTSP) and DVB-H channels.

3.3 IPTV Platform

Unlike other access platforms, all IPTV information (content, billing, interactivity) is sent to the user over the same connection. Since it is based on the TCP/IP architecture, general content is being transmitted to everyone (Scheduled Content) while personalized and specific content can be selected and transmitted to each user (Content On Demand, CoD), so the user becomes the actual administrator of the TV contents.

IPTV services are offered through specific receiver devices connected to the TV (Set-Top Box) or televisions with this technology incorporated. Thanks to the increased capabilities in terminals, it is now possible on IPTV to unify the technologies employed to represent menus on screen and interact with the user by means of Web solutions, in such a way that programming languages as Java, Javascript, CSS and XML (AJAX) are used, so dynamic content visualization could be possible.

The user interface has especially been designed for the TV environment, adapting graphical components and the interaction is carried out through a remote control.

In order to transmit multimedia content through the IP network, video content is encoded using MPEG-2 or MPEG-4 (H.264), audio is encoded with MP3 or AAC coder, and they are encapsulated into a transport stream or container (MPEG-TS). Then, the resulting stream is sent over RTP/UDP (or UDP only) to a Multicast address for live TV. Streams can be transmitted over RTP/UDP or HTTP/TCP over Unicast IP in case of video on demand.

3.4 Kiosk Platform

We consider the kiosk as a public access point to RAUDOS Platform. This access point is based on a physical infrastructure, based on a computer with a touch screen and Internet connection. Additionally, the computer has a RFID reader and a Bluetooth device. The objective of these two elements is to provide, respectively, the identification process in a natural and easy way (RFID), and to allow download to mobile of different contents, previously seen in the kiosk and in an adequate format for the mobile phone.

The development of this application has been performed in .NET environment and it has different screens, with specific functionality: interaction with RFID tag and virtual keyboard, video and activities visualization and Bluetooth mobile download.

Kiosk turns into the ideal platform, not only for enjoying audiovisual content in the specific place, but also for obtaining those contents for other occasions. Additionally, it is a strategic access point for recruiting new users for the platform and it can work as a potential audience recruiter.

4. RESIDENTIAL GATEWAY

This section focuses on the task of finding and processing multimedia contents in the residential environment. Such contents can be either the user's local ones or RAUDOS contents, located on a remote server. An extension to the UPnP AV technology has been developed that can manage both types of content in a way that is homogeneous and transparent to the user.

The UPnP generic architecture is based on two types of elements: the Devices and the Control Points. The devices offer services to a network and the Control Points search, register and use them. On top of this generic architecture, other extensions for specific devices in different application domains are defined. This development extends the AudioVisual UPnP standard (UPnP AV V1.0) ([2], [3]). This standard provides a mechanism for sharing multimedia contents among different types of devices with different formats and transfer protocols.

Two AV specific devices (a MediaServer and a MediaRenderer) and an AV Control Point have been implemented as OSGi services [4]. They extend the basic functionality to allow discovering remote RAUDOS contents and interacting with them.

5. PLATFORM SELECTION

RAUDOS authentication mechanism enables user identification

and authorization, providing continuity while accessing separate and isolated services. In such a highly heterogeneous environment, the best approach requires each application to be responsible for credential collection. Credentials are transferred to the User Manager that performs the validation procedure, providing the requester Application the authorization state and profile of the User. The User Manager implements an SSO (*Single Sign On*) mechanism where each successful authentication generates a temporal *Service Token* that Applications use to provide continuity through services.

RAUDOS must also provide a single *Service Entry Point* able to adapt Application to client presentation capabilities. A User Agent based Proxy is implemented. Each individual request is matched against an internal Service point based in the User Agent. This philosophy enables an orthogonal integration of services where JavaScript or Flash portals can share a space with XHTML only servers. The SoTA analysis shows there is any proxy implementation supporting this feature, so it is implemented one based in the WURLF [8] library.

6. INTERACTIVE ACTIVITIES

Platform's interactivity is achieved mainly through the Interactive Activities, which serve as an element for the creation of customer loyalty as well as advertising media for the advertiser agents. The activities are published in online web format, and are accessible from browsers available for every access technology. This fact allows designing applications with a common *Look&Feel* for easy distribution. To start with, we have developed two activities consisting of a quiz contest and a survey application.

Customers participating in the quiz contest can win virtual credit by answering correctly a number of questions with an incremental difficulty level. This credit can then be used for buying multimedia content within the RAUDOS Platform.

On the other hand, the survey application provides a thematic entertainment depending on the approach you take to the activity. It may have a special interest for enterprises, stores or advertising agents, as they can get useful reports about current and future products released to the market in return for some virtual credit given to platform users.

7. CONTENT MANAGEMENT

The Content Manager is responsible for the storage of content and metadata in the respective repositories in order to be able to be edited, deleted and/or displayed.

Contents, which are uploaded to the content server, are stored for subsequent automatic transcoding for adaption to different platforms. Their corresponding metadata are stored as well.

The metadata are defined and organized according to the MPEG-7 and MPEG-21 standards [8][9]. MPEG-21 offers a standard system of organization and encapsulation of the metadata contained in *Digital Items*. MPEG-7 gives us the definition of metadata, both semantic (title, genre, summary, developers...) and technical (duration, bit-rate, format, size ...).

Some metadata are created specifically for the RAUDOS system, such as raudollar indicating the cost of content; the "value of recommendation" used for display and for voting by the users; the "thumbnail" which shows a small picture of content; and content type (video or TV channel).

The list of genres and subgenres, which categorizes the content in the manager, is based on the TV-Anytime standard [10], from which the most representative subset for cataloguing the contents has been chosen.

The content search is made by specifying search criteria as well as a criterion for sorting the results. Search criteria are based on the definition of metadata, with the possibility to create ranges or limits. The sorting allows us to quantify results and sort them by field.

Finally, the content manager provides a URL for customers to access the view of content from a streaming server, according to the RTSP protocol.

8. CONTENT RECOMMENDATION

The recommendation system of RAUDOS uses the user profile in order to recommend multimedia content as well as banners. This profile is created from user data (age, sex, study level, etc.) and from data which are inferred from the user behaviour through the Platform (clicks, ratings, etc).

The *cold start* problem [5] constitutes one of the main problems of the recommendation systems. In order to solve this problem, we have used the *Ontology Filtering* (OF) technique [6] which uses categories to infer the ratings of an item given by a user. These categories are structured in a tree which can be built manually or automatically by means of different *clustering* techniques.

The selection of the appropriate tree is made by dividing the user ratings between two sets, training and test sets. Then, different clusters are generated from the training subset and the test subset is used to check which cluster technique produces the best result. In this way, different users can use different clustering strategies.

To verify the quality of the obtained results using ontology filtering we have used 100,000 ratings from 943 users and 1,682 movies from the Movielens dataset [7]. The results are showed in the following figure:

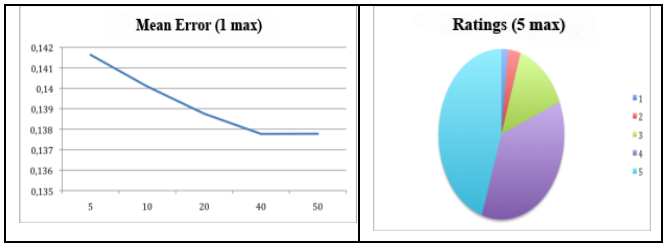


Figure 2: Recommendation Performance Analysis

The left figure shows the medium average error considering one as maximum error and the number of items which have been rated by the user. In this figure the error is under 0.142 (using 5 stars rating mechanism is always under 0.7), which is a good result considering that only five items have been rated.

The right figure shows the items which have been recommended according to the user ratings. In this sense we can observe that 80 percent of items have been rated with 4 and 5, and only 4 percent of items have been rated with 1 and 2, that is, the probability of getting a bad recommendation is about 4 percent.

9. VIRTUAL CHARACTERS

Up to date, user interfaces are based on WIMP paradigm (Windows, Icons, Menus, Pointer). Nevertheless, this interaction paradigm limits the number of channels that persons use for communication, mainly due to two reasons: sensorial skills of persons are ignored and interfaces are far away from natural interaction between persons. Moreover, WIMP interfaces do not fit the interactive digital television paradigm [11].

In the RAUDOS project, in addition to the development of a traditional interface, an interface based on virtual characters has been implemented. The purpose of this interface is to do the interaction more natural and intuitive [12]. Virtual characters are defined by an appearance and behaviour, and they try to simulate a natural communication between the user and the system with the help of different technologies as Text To Speech (TTS) or corporal animation.

The availability of these two paradigms for different platforms developed in the RAUDOS project will allow further research on the advantages and drawbacks of each model, as well as about user preferences.

The avatar used in the project is based on a 3D model. Due to the lack of support to render computer graphics in most of the platforms targeted in the project [13], we have adopted the solution of rendering the content in a server and delivering the video using streaming technologies [14].

In the implementation of the Virtual Characters Manager, we use the *gstreamer* RTSP server. To do the actual rendering of the avatar, we have developed a *gstreamer* plug-in. For each client that access an application, a new pipeline is created; this pipeline streams the rendered video of the avatar to the client using standard RTP protocol. The interaction of the user with the avatar is done by web services. The Virtual Characters Manager instructs the corresponding pipeline to update the behaviour of the avatar. The Virtual Character Manager is also responsible of encoding the rendered video with a resolution and a codec that best fits with the final device.

10. FUTURE WORK

RAUDOS future work will be centred in four main aspects: access, delivery, content, and validation.

Although RAUDOS implements access with the main broadcast technologies combined with unicast ones, it is possible to add some other technologies like DVB-C and DVB-S.

RAUDOS delivery could be optimized in order to achieve greater scalability using techniques like P2P streaming through Internet or Quality of Service (QoS) mechanisms in ISP, like intelligent routing.

Content control in RAUDOS system is made in a centralized way in order to ensure only professional distributors can put contents in the system. The next step is developing new multiplatform tools in order to put user and free Internet content into the system.

Finally, RAUDOS system must be validated with end users in a real context using Living Lab methodologies [15].

11. ACKNOWLEDGMENTS

This research has been funded by Ministry of Industry, Tourism and Trade and European Regional Development Fund under the RAUDOS project.

12. REFERENCES

[1] X-Smiles homepage: http://www.xsmiles.org/

[2] UPnP AV Architecture: 1, http://www.upnp.org/specs/av/upnp-av-avarchitecture-v1-20020622.pdf, 2002

[3] D. O. Kang, K. Kang, S. Choi, J. Lee "UPnP AV Architectural Multimedia System with a Home Gateway Powered by the OSGi Platform", IEEE Transactions on Consumer Electronics, February 2005

[4] Open Service Gateway Initiative (OSGi) Alliance, URL: http://www.osgi.org, May 2009

[5] A. Schein, A. Popescul, L.Ungar, and D. Pennock. Methods and metrics for cold-start recommendations. In *25th* Int. ACM SIGIR'02, 2002.

[6] V. Schickel-Zuber and B. Faltings. Using Hierarchical Clustering for Learning the Ontologies used in Recommendations Systems. ACM *KDD'07.*

[7] *Dataset de Movielens* http://www.cs.umn.edu/Research/GroupLens/data.

[8] ISO/IEC JTC 1/SC 29 15938. "MPEG-7 (Multimedia content description interface)"

[9] ISO/IEC 21000 - Multimedia framework (MPEG-21)

[10] Metadata Specification of TV-Anytime, http://www.tv-anytime.org/workinggroups/wgmd.html

WURLF http:// wurfl.sourceforge.net

[11] Pablo Cesar and Konstantinos Chorianopoulos (2009) "The Evolution of TV Systems, Content, and Users Toward Interactivity", Foundations and Trends® in Human–Computer Interaction: http:/dx.doi.org/10.1561/1100000008

[12] Ortiz, A., Oyarzun, D., Carretero, M., and Garay-Vitoria, N. (2006),"Virtual characters as emotional interaction element in the user interfaces", AMDO 2006. IV Conference on Articulated Motion and Deformable Objects. Andrax, Mallorca, Spain.

[13] A. Ugarte, I. García, A. Ortiz, and D. Oyarzun, "User interfaces based on 3D avatars for interactive television," in European Conference on Interactive Television, 2007.

[14] C. Nguyen, M. Tuan Le, D. Yoon and H. Kim, "A Graphics Adaptation Framework and Video Streaming Technique for 3D Scene Representation and Interaction on Mobile Devices", 4th International Conference on Universal Access in Human-Computer Interaction, UAHCI, 2007.

[15] European Network of Excellence of Living Labs http://www.openlivinglabs.eu/

MORE AUTHORS: J.V. Llácer (jvllacer@aido.es), J.Agreda (jagreda@euve.org), E. Loizaga(eloizaga@euve.org), J.C.Guerri(jcguerri@dcom.upv.es), P.Arce (paarvi@iteam.upv.es), J.M.Quinteiro(jquintei@iuma.ulpgc.es), E. Martel (emartel@iuma.ulpgc.es), N. Martínez (nati@it.uc3m.es), F.Pinyol (fpinyol@salle.url.edu), J.Martínez (jmf@inv.it.uc3m.es), J.M.Tejada (jorgetejada@det.uvigo.es), D.González (dagomar@enigma.det.uvigo.es), F.J.López (fjlopez@solaiemes.com), C.Galicia (cgalicia@rtvv.es), V.Beviá (vbevia@rtvv.es), M.Zorrilla (mzorrilla@vicomtech.org), A. Arrieta (aarrieta@vicomtech.org)

DVB-PCF to Flash Lite Transcoder

Ramon Martín de Pozuelo
GTM - Grup de Recerca en Tecnologies Mèdia
LA SALLE - UNIVERSITAT RAMON LLULL
Quatre Camins 2, 08022
Barcelona, Catalonia, Spain
ramonmdp@salle.url.edu

Francesc Enrich
GTM - Grup de Recerca en Tecnologies Mèdia
LA SALLE - UNIVERSITAT RAMON LLULL
Quatre Camins 2, 08022
Barcelona, Catalonia, Spain
frane@salle.url.edu

Gabriel Fernández
GTM - Grup de Recerca en Tecnologies Mèdia
LA SALLE - UNIVERSITAT RAMON LLULL
Quatre Camins 2, 08022
Barcelona, Catalonia, Spain
gabrielf@salle.url.edu

ABSTRACT
The heterogeneity of tools in interactive television applications development makes more difficult a general deployment of them. In order to solve that, DVB-PCF (Digital Video Broadcasting-Portable Content Format) is a standard that aims to consolidate a standard way to define these applications, allowing an easier interchange of them, regardless of their platform target.

However, in order to support this standard, it will be necessary to create converters that could interpret its descriptions and create the platform-specific applications. In this context, this paper introduces the design, implementation and results of a DVB-PCF to Adobe Flash Lite transcoder, which intends to render a DVB-PCF description, as accurate as possible, in an Adobe Flash Lite 3.1 application.

Categories and Subject Descriptors
J.m [**Computer Applications**]: Miscellaneous;
I.7 [**Document and Text Processing**]: Standards

General Terms
Algorithms, Design, Standardization, Languages.

Keywords
DVB-PCF, Flash-Lite, Interactive TV, transcoder, parser, converter.

1. INTRODUCTION
Currently, the increasing market of digital television and interactive applications for non-PC devices does not have a consensus standard that guides how to develop this kind of applications for set-top boxes, hand-held devices, etc. The variety of mediums and platforms where those applications are deployed complicates the development of a proposition to be delivered to multiple target platforms, due to physical, operational and commercial constraints [4]. This fact brings us to a situation where all the developers make their applications for a specific target platform, and it is hard to create a broad market around interactive digital television (iTV) services.

However, looking at this panorama, in 2006 emerges DVB-PCF (Portable Content Format) [1]. The PCF is a standard focused in the description of interactive services by describing the visual experience to be shown to the user. It provides the industry with a platform-independent format for the business-to-business interchange of interactive content, and consequently a means for increasing the interoperability of authoring tools, head-end systems and broadcast networks.

Using PCF, content developers will deliver PCF content to network operators or broadcasters, who will then run it through a translator to create something appropriate for their network. It is not intended to be transmitted to final users' set-top boxes or integrated digital TVs. It enhances interoperability by enabling content providers to author their content once and run it on multiple API platforms, covering 80% of interactive television applications [5]. This opens a market that has been completely diversified and hard to exploit efficiently, but PCF is not an end-to-end system that could solve this immediately. Indeed, although it becomes the followed standard to define most of iTV services, other technologies are needed in order to adapt PCF content to every platform and finish the distribution chain.

The recently supported for many Digital Home devices [2] [6] [3], Adobe Flash Lite presents a powerful tool that paves the way for rich Flash-based entertainment experiences on televisions that offer viewers new options for accessing web content on their TVs, phones, etc.

Taking these two facts into account, DVB-PCF to Flash Lite Transcoder wants to fill the gap between the two technologies and bring a solution that can interpret a PCF description and render an application that could be visualized in any device that supports Flash Lite Player.

2. TRANSCODER FRAMEWORK

2.1 Development process

The development of this transcoder has been performed in two phases. In the first one, a DVB-PCF interpreter was written in ActionScript (AS) 2.0. As a result of this, a Flash Lite application that dynamically renders a DVB-PCF description was obtained. It can render a service, scenes and basic elements as images or textboxes, and a simple navigation between scenes using basic menus. This resultant application should always be accompanied by an XML file with the DVB-PCF description, and it will render it when loaded. It is a relevant fact, because it permits to dynamically change the application by modifying the description and reloading the application. For this purpose, data model of basic DVB-PCF elements was written in AS 2.0, defining a set of basic PCF elements. However, as interesting as it could be to modify dynamically the rendered application, it sometimes falls back to security issues and would not be useful in some environments.

In a second phase, a complete transcoder was made, which permits the creation of iTV services without the necessity of the description, once the application had been created. This offers more flexibility in some aspects, as the service provider does not need to send multiple files to its users: just a standalone application is supplied. It also solves the security problem posed by the fact of delivering an application that could be modified by changing the associated XML description file.

Taking into account the current economic model of iTV, it will not be interesting for service providers to give their users the possibility of modifying the application by themselves. It would be like providing all the source code to their clients, allowing them to modify it without restrictions. It would be interesting to give users the chance to create new features by simply changing an XML file, but it could be the end of iTV market as we know it.

As a result of this development, the new transcoder only needs the description and images of the service in compilation time, when it creates an Adobe Flash SWF file, a self-contained application that could not be modified, and a easier way to deliver the application to users.

2.2 Package structure

This section details the structure of the final transcoder mentioned above, and how it was designed. First of all, different modules were established, foreseeing the possibility to swap some of them to other technologies, especially those that are more platform-specific, like the data model.

The packages finally defined were the following (also shown in Figure 1):

- **DVB-PCF ActionScript Data Model** package includes all classes related to AS objects that can form a PCF description.
- **LoadManager** package contains all classes related to load PCF information into memory including image loading issues.

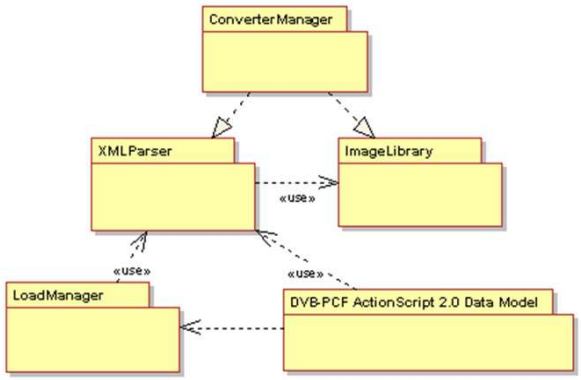


Figure 1: DVB-PCF to Flash Lite transcoder packages

- **XMLParser** package includes all classes related to PCF description interpretation, and use LoadManager and Data Model package to create an AS file that includes all the elements defined in PCF description.
- **ImageLibrary** package classes control image imports and create a library with all the images specified in the PCF description file.
- **ConverterManager** package includes all classes related to complete the conversion and make up the SWF file, managing the outputs from XMLParser and ImageLibrary.

First two packages were written completely in ActionScript 2.0, whereas the following three were written in Java.

Trying to provide a way to adapt the description of the service as exact as possible to the final application rendered, the first step in design stage was an extensive study of all the elements defined in DVB-PCF standard and how they could be mapped into Flash parameters.

PCF should basically provide means to allow application authors to describe the intended experience, so most of its elements are fundamentally visual. Flash Lite 3 uses AS 2.0, and it is a limited platform in comparison to Flash Player used in higher power devices, but it still has enough potential to make it feasible to recreate all PCF basic features by combining AS 2.0 components.

Resulting from that study, a package of AS classes describing most DVB-PCF elements is created. This package defines PCF elements (Service, Scene, Background, Image, Menu, etc.), providing its required parameters to each item, so that it could be represented as exact as possible to what is described. They also provide getter and setter functions for each parameter to easily interact with them, and a basic constructor that only initializes all the parameters of the class.

These classes do not provide methods to render all these features in Flash; their intention is only to work as containers, maintaining the PCF element information.

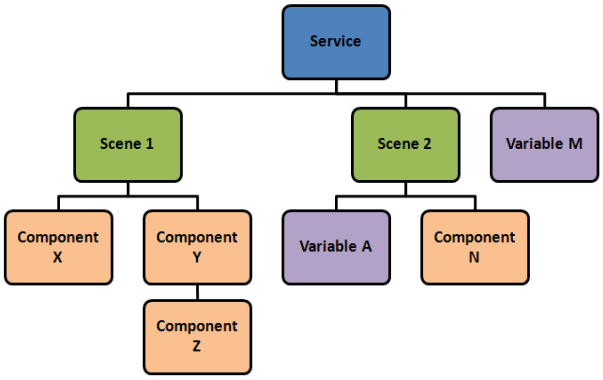


Figure 2: Hierarchical tree sample of DVB-PCF Components

Once this data model had been accomplished, and it made possible to storage every feature of PCF elements, a render package was needed. This package will be the responsible of taking all the parameters stored in every element defined and render it as exact as possible.

Two more packages were defined in order to control the load of all the information in compilation time: (1) a set of classes to load the images included in DVB-PCF description, and (2) a parser that reads from DVB-PCF description and creates AS 2.0 files. These resulting files, in addition to `Data Model` and `LoadManager`, are used to compile the final application.

2.2.1 DVB-PCF Data Model

PCF service descriptions are formed by hierarchical structures of information items, most of them represented as PCF components, and each one with a unique name identifier. Figure 2 shows an example tree of its typical structure.

The root element of a description is `service` item, which has to contain one or more `scene` items, and define which of these scenes will be the first to be shown to the user. An important fact to emphasize is that the service should never show more than one scene simultaneously in its life-cycle. All the scenes that could be rendered have to be inside `service` item or some of its sub-containers.

As shown in [4], main aspects of a basic PCF service description could be summarized in:

- **Components :** data blocks that build the description.
- **Content :** managed and presented using PCF Components.
- **Behaviour :** responsible of the generation and control of events generated in execution time.

To support these aspects, the data model of the transcoder has to contain a minimum set of DVB-PCF elements to create a real iTV service experience. This set includes visual components (`Scene, Images, Textbox, Rectangle, etc`), content components (`Size, Position, Color, URI, etc`), behaviour components (`OnEvent, SceneNavigate, Target`) and others that involve visual aspects, but are also closely related to the control of the service (`Service, Scene, Menu`)

2.2.2 Load Manager

The way that all objects are rendered in the application is independent from how they are created, so it implies specific load functions for every object. `LoadManager` package contains a set of classes, almost one for every PCF element supported in the transcoder. Each of these classes obtains all the features information of the element and creates the visual equivalent object in Flash. When the complexity of DVB-PCF elements included increases, a bigger combination of ActionScript objects is usually required, in order to be more accurate. But the chosen method has been effective up to now. This package, jointly with DVB-PCF Data Model, was exactly the same in the first version of the interpreter, mentioned above.

2.2.3 Description parser

This module is completely implemented in Java. It makes XML reading easier, permitting to recognize every PCF element and creating an AS file. Based on previous `Data Model` and `LoadManager`, these files contain only simple calls to these modules, first initialising its elements and then calling its specific loading function. This separation is what permits to change AS 2.0 modules by other technology ones, reusing Java modules to create another transcoder.

2.3 Scene navigation

The most difficult task of generating a iTV service using an XML description is to store all the interactivity information. In a standard Flash application, it is very easy to define a button or a menu that makes a specific task. In the case of the transcoder, it has to define these actions and handle them in a way as generic as possible, because the application controls are created on the fly, and associated to elements, as the application is created sequentially. It requires managing and creating a context for the application, and monitor it, as a button could do different things depending on the life-cycle point of the service.

In addition, DVB-PCF specification has precisely defined the way of how to define events, and it was very helpful in the definition of the navigation between scenes, but there are implicit events that are more undetermined, like the navigation within a scene. However, this transcoder implements a simple navigation between all focusable objects that are in a scene using *arrow keys* of the device or remote control.

2.4 Development tools

The development of this transcoder is completely multiplatform and open source. It has been implemented using Java and Open Flash tools, as Mtasc and Swfmill compilers, over Eclipse IDE framework. It offers a flexibility extra point, so transcoder can work (and has been tested) over different operating systems like Windows, Linux, etc.

The Flash Lite applications resulted from this transcoder has been tested using Flash Lite 3.1 Player in different devices (a Pentium IV 2.2GHz PC and different types of set-

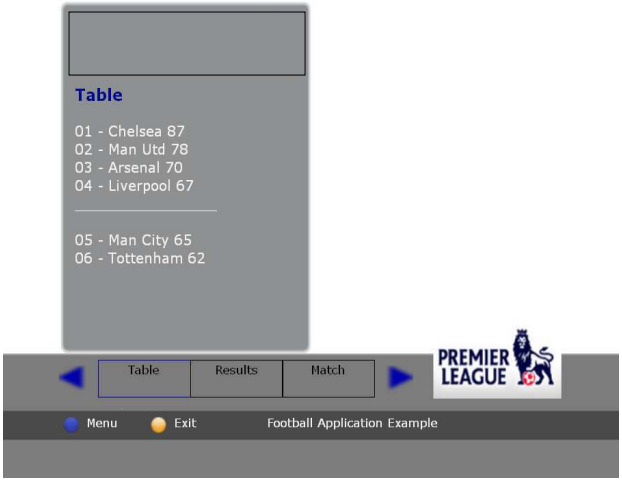


Figure 3: Football Interactive Application Sample

Figure 4: Multiple Menu Interactive Application Sample

top boxes), with good processing performance in all of them. In its current version, transcoder can efficiently render most of the visual effects presented in DVB-PCF. In order to test these elements and to check that every feature is depicted correctly, some interactive services were designed.

3. RESULTS

As an example of interactive test service, a football interactive application is devised in order to be sent during a football match broadcasting. This application shows different scenes where user can navigate using their remote control's left and right arrow keys. Figure 3 shows a scene of this application that contains the league's classification table, and a horizontal menu that permits the scene navigation control. In addition, in this scene there are three different focusable objects (horizontal menu, a button to go to main menu and an exit button) that user can focus using up and down arrow keys. In case that the scene contains vertical menus, the control will swap, moving up and down within the menus, and left and right to move between objects.

This feature is also shown in Figure 4, where an application with a double menu scene is presented. User can move from one menu to another horizontally and vertically within them. This method of handling navigation in the applications rendered by the transcoder, limits the use of horizontal and vertical menus in the same scene, so it still needs to be improved in next versions.

4. CONCLUSIONS

In this paper, a solution for a transcoding application from DVB-PCF to Flash Lite is presented, allowing finalizing the iTV service production system proposed by DVB-PCF standard. This application has proved the potential of DVB-PCF format and it has presented how is possible to render DVB-PCF interactive services descriptions using Flash Lite, a widely used and on its way up format.

In this sense, it is important to emphasize the simplicity of creating applications using descriptive XML files, and transforming them into real applications. It brings forward the possibility to create interactive services even to people that do not have any technical knowledge. This method also permits to create applications regardless of the final target, and reuse their descriptions integrally (or only part of them) to easily generate new services.

DVB-PCF elements supported by the first version of the transcoder, as shown in the results section, are enough to generate a powerful interactive experience to user, showing many different types of visual objects. It also provides support to some event elements defined in DVB-PCF specification that offers to the user a complete and intuitive navigation among different objects and scenes of the service. However, this version still does not support more advanced descriptions that allow videos, other more complex graphical objects or return channel connections. These features are still in development and they are expected in next versions.

5. REFERENCES

[1] Digital video broadcasting; portable content format (pcf) specification 1.0. ETSI TS 102 523 V1.1.1, 2006.

[2] Adobe. Adobe and broadcom bring the adobe flash platform to tvs. 2009 International CES, press release, January 2009.

[3] BBC. Adobe flash secures set-top deal. http://news.bbc.co.uk/2/hi/8008070.stm, April 2009.

[4] R. Bradbury, R. Cartwright, J. Hunter, S. Perrott, and J. Rosser. Portable content format: a standard for describing an interactive digital television service. R&d white paper whp 134, BBC Research & Development, 2006.

[5] P. Bugnot. Commission of european communities; commission staff working paper on the interoperability of digital interactive television services pursuant to communication com(2004), 2004.

[6] E. Eldrom, S. Janousek, and T. Joos. *AdvancED Flash on Devices*. Apress and friends of ED books, 2009.

Using video and social networks to improve the access of daycare centers to museums: one case study

Diana de Sousa
Media Lab, Aalto University School of Art and Design
Hämeentie 135C
00560 Helsinki Finland
+358 45 275 2677
dspsousa@gmail.com

Mariana Salgado
Media Lab, Aalto University School of Art and Design
Hämeentie 135C
00560 Helsinki Finland
+358 40 585 7727
mariana.salgado@gmail.com

ABSTRACT

In this paper we address how social media tools can improve the access of daycares to museums. Our proposal is based on ongoing case study started at the Design Museum Helsinki in 2007. This case study is *Oswald and the Objects*. It aims at making the museum collection accessible to daycares that are located far from the museum. We describe the case study and reflect on the possibilities of social media tools that can be used by a young audience to make the museum more accessible and engaging. Using and designing the implementation of social media tools has made us aware of the need to plan particularly carefully due to sensitive privacy concerns and technical reliability issues. We explain our reasoning behind the selection of services for sharing photos, documents and videos. We try to show how the specific needs of the project have guided our choice of one service over another. In parallel to this discussion we try to analyze the possibilities of using video for the creation of interactive Web TV in the context of museums.

Categories and Subject Descriptors

J.5 [**Arts and Humanities**]: *Fine Arts*

General Terms

Design, Experimentation.

Keywords

Design education, informal learning, video, digital media, social media, social networks, user-generated content, internet, daycare, museum, accessibility.

1. INTRODUCTION

Design Museum Helsinki [1] currently offers activities for children but they happen mostly at the museum's premises. This is a major restriction for daycares located far from the museum, since it's difficult to take small children on long distance journeys. As an attempt to address this problem five workshops were organized at a daycare in Helsinki in 2007. They were the *Oswald and the Objects* workshops (Oswald is the name of a puppet), part of the project *The Secret Life of Objects* [2].

Our current project, also named *Oswald and the Objects* (in Finnish: *Esa ja Esineet*) consists of developing an online service and a web platform for improving the access of Finnish daycare centers to educational activities offered by the Design Museum in Helsinki. The project involves collaborators from four different entities: Design Museum, Media Lab Helsinki, Zipipop and ConnectedDay. At present the project is in the prototype stage. A team of museum docents have been creating content for the web service and a team of developers have been building the website. A round of tests at daycares is scheduled for March 2010.

The service targets daycare educators and children from 3 to 6 years old. It provides guidelines for the educators on how to organize activities at daycares related to the museum's collection. It also offers some of the necessary educational resources, like videos and printable presentations to conduct autonomous activities related to the museum collection. Our aim is to improve the engagement of daycare educators and children with these activities by using social media tools to share materials (photos and videos of the children's works) among the daycares and the museum. In particular, we discuss video sharing as an opportunity to start a museum Web TV directed at daycares.

In this paper, we analyze the case study *Oswald and the Objects* to explore answers to our research questions: how to design and implement social media tools in the context of museums? How can shared video content be reinterpreted as a museum Web TV channel?

2. THE SERVICE IN BRIEF

Back in 2007 the workshops were led by the members of the museum's education team and included a live puppet show performance using the puppet Oswald, who introduced the children to a different design object in each workshop. A dialogue engaged the puppet and the children.

With the online version of *Oswald and the Objects* [3] we try to offer children an experience similar to the one they would have if the puppet Oswald and the museum docents could visit them in person. Our purpose was not to simply translate the *Oswald and the Objects* workshops to an online environment, rather enhance it with the possibilities of the online media and tools. This "blended" approach aims towards blurring the on-site activities with the online interaction by using social media tools for sharing materials (photos, videos, audio and texts). In this informal learning environment children are learning both from the resources offered by the museum and from their peers, with the materials shared by other daycares. The process of contributing materials, together with commenting and replies from all parties, should create an open dialogue between the daycares and the museum.

Before starting the activities, the children watch a video of the puppet Oswald from the *Oswald and the Objects*'s website. This replaces the live puppet show. Afterwards follows a discussion about the design object in the context of the children's daily

experiences, moderated by the daycare educators (instead of the museum docents). Subsequently, the daycare educators lead the children through a series of creative activities where they draw, model or design new objects or new contexts for the objects.

We encourage daycare educators and children to take pictures or make short videos of the works and processes to later share online. They can write comments to those images and videos. For example, the children can tell short stories about their own work. If the child cannot read or write the educators can help out.

The comments are a conversation starting point between the parties. The museum docents send a reply to the children or the daycare by also adding the puppet Oswald's comments to the photos or by publishing a short video where Oswald addresses and thanks the children for their works. It's also possible for different daycares to share materials and let the children have discussions about them.

3. IT IS ALL ABOUT SHARING AND DISCUSSING

The online version of *Oswald and the Objects* combines two different web services: the public website *Oswald and the Objects* [3], owned by Design Museum, and the online community website of ConnectedDay [4]. The main language is Finnish.

Figure 1. The *Oswald and the Objects* website.

The *Oswald and the Objects* website is used for delivering the necessary information, documentation and materials about the educational activities from the museum to daycare educators. The materials include videos, photos, guidelines in Powerpoint, PDF or slideshow formats. There is also a small discussion forum for enabling communication between daycare educators and the docents in Design Museum. The forum is equally open to the public for collection of comments, questions and suggestions.

3.1 Discussing

We planned to enable discussion forum features in some of the pages of *Oswald and the Objects* website. The main purpose of these forums is to collect the experiences, concerns and suggestions of the educators about the educational activities. The discussion forums are addressed to a set of activities and the themes of these activities. Each set of activities has its own webpage and its own discussion thread. We made the discussion threads visible in several places on the website to increase the probability of people contributing to them. In this way we give more visibility to the threads and encourage active participation from the audience.

Another approach to encourage contributions is to pre-populate the discussion threads with comments, questions and answers before launching the website. This initial content has the purpose of breaking the ice and giving leads concerning what kind of feedback the Design Museum would like to receive.

Lowering the barriers to people's interaction is also important, as we learned from a case study about visitor participation at the Delaware Art Museum [5]. Thus in *Oswald and the Objects* there is no need to register or log-in in order to comment in the discussion forums. They are public and everyone is welcome to contribute or send feedback. We are aware that this may also allow for a number of unwanted or inappropriate comments, but the website administrator will get an email notification for each new post, which will help with the monitoring of the forums. This post moderation of the comments after online publishing has already proved successful in other museum projects [6]

3.2 Sharing documents

Design Museum needed to share PDF and Powerpoint files with the daycare educators in order to offer them more specific information about design history and point out the most important themes to be discussed during the educational activities. We decided that the simplest way to share these documents would be to use an external service for document sharing that was reliable, free of charge and which gave a possibility to embed the documents in *Oswald and the Objects* website. We considered the embedding option to be the decisive factor for choosing to use such a service, since this allows the visitors to browse inside the document directly from the website, without needing to download it beforehand.

We looked at the two main options among online document sharing services: Scribd [7] and Slideshare [8]. They both offered very similar features, so we chose the one that had the most visitors' traffic — in the assumption that the more traffic a service has the more established it is and the less likely it is to be discontinued in the near future.

However, when we tried to embed the document in our website we found that the Scribd player hides the control tools from immediate view. Slideshare was different. Their player shows the most important control tools at the bottom, like most video players. There is a moving arrow floating on top that invites users to click for the next page. Additionally, it allows direct sharing of documents to Facebook, Del.icio.us and a number of other popular social networks. One can also send the document by email to a friend directly from *Oswald and the Objects* website just by using the Slideshare tools. Slideshare has also a more appealing visual interface. Because of all these reasons we finally decided to choose Slideshare to share documents.

3.3 Sharing materials from daycares

We needed a channel with restricted access where daycares could share photos, videos, stories and comments from children about their works and receive feedback, comments and replies from the docents in the museum. We initially planned to use Flickr [9], since it can handle both photos and videos and it has privacy options.

We changed opinion when we learned that ConnectedDay already had its own secure system of photo and video sharing especially aimed at daycares. ConnectedDay is a community site aimed at professional daycare providers, with the goal of increasing parental involvement. It provides a secure picture and video sharing service from daycares to parents. ConnectedDay had around 150 daycare and preschools using their service in Finland as of mid 2008 when we asked them to collaborate in our project. They agreed to participate by providing one camera phone and one account in their service for the Design Museum, free of charge for one year.

When the daycare educators upload the children's photos, videos or stories to ConnectedDay they can share them with the museum. The museum can then access the images and comment on them as well; through ConnectedDay the museum can also send a short video of the puppet Oswald thanking children for sharing their works. In addition, the museum can act as a dialogue moderator by putting different daycares in contact with each other, sharing images from one daycare to another; the children will be able to see what kind of works and ideas came from their peers in different cities and send their comments to them. These are possibilities that we plan to test in the future.

The idea of using other photo and video sharing services for materials shared by the daycares was not completely been put aside, since ConnectedDay is a paid service and many daycares are yet not using it. The museum docents wish to make *Oswald and the Objects* accessible to the largest number of daycares, therefore a solution for the ones which are left out of ConnectedDay's network will be prepared in the future. But for the moment our prototype uses this service only.

3.4 Sharing video from Design Museum

The videos provided by the Design Museum are currently being shared through three different platforms: Vimeo [10], YouTube [11] and ConnectedDay [4]. Cost efficiency was the principal reason for choosing Vimeo and YouTube video sharing services. They are both ready to use and available for free. This way our scarce resources could be directed to producing the actual contents instead of buying media space. Also most video sharing services already include several social interaction tools which can benefit our project. For example:

1. Commenting and rating options, which can be used as an extra tool to collect feedback from a wider audience than just daycare educators.

2. Tagging options, which improve the chances of videos being found by people searching for related topics or keywords, and thus help spread awareness of the project.

3. Direct sharing connections to other social networks, allowing everyone to share the videos and publicize the project easily.

4. Statistics and view counters, which can give us a good idea of how popular the videos are, where our audience is located and where the website is getting traffic from.

We decided to use Vimeo and YouTube simultaneously for precaution. The services may suddenly cease to exist and there is no guarantee that they will be free of charge forever. Thus, if one service becomes unavailable we can alternatively use the videos hosted in the other service. Using several services in parallel also brings the benefit of increasing the possibility of people finding and sharing the videos, giving more visibility to the whole project.

The reason to use ConnectedDay had mainly to do with logistics. ConnectedDay provides daycares with camera phones and training in using those camera phones to collect photos and videos (which are directly uploaded to the web from the phones). They also train educators to upload and share the contents in ConnectedDay's secure web environment. They have a structured service already in place, which otherwise would have to be built by the museum.

3.5 The purpose of video contents

Video is being used in several ways and for different purposes at different stages of the activities proposed in *Oswald and the Objects*.

3.5.1 Video-samples of the activities

These videos show the work-flow and outcomes of the workshops done in 2007. They were filmed during the events at the pilot daycare. The purpose of these videos is to instruct and encourage the educators to use the service. As we have learned from the series Design Squad [12] and DragonflyTV [13], the videos show a step-by-step sequence of the activities and exemplify the results achieved by the children. These videos also show how much fun the children and educators can have while doing the proposed tasks. Their role is to stimulate educators to engage with *Oswald and the Objects*. They are an important activating factor.

3.5.2 Videos of puppet Oswald and the theme object

These are videos of the puppet Oswald and his stories about the themed design object for each set of activities. They are viewed by the children before they start each activity. They introduce the activity's theme (as for example the Aalto vase, Teema mug). The videos stimulate the children's analytical point of view by presenting the object out of its usual everyday context.

3.5.3 Videos made by the daycares

These videos document the works and processes developed by the children during the *Oswald and the Objects* activities and they are made by the daycare educators.

They are located at ConnectedDay's website and can only be viewed by people authorized by the daycare educators. These videos allow docents in the Design Museum to follow the results achieved by the children during the activities and give feedback to children about their works. If they wish daycare educators can share the videos with other daycares to compare experiences.

3.5.4 Videos of Oswald's replies

These are videos of the puppet Oswald addressing the children about their work. When the daycare educators share images of the children's works with the museum, the museum docents send a video of the puppet Oswald back to the daycare thanking the children for sharing their works. This video is pre-recorded and the museum shares it with each daycare individually through ConnectedDay's website. At the moment the video is the same for all daycares but in the future it may include personal messages, for example by mentioning the names of the addressed children

and making specific comments on their work.

4. CONCLUSIONS

In this case study, we encourage daycares to connect with the museum collection, the museum docents and with other daycares by using video, photo and document sharing services, and online discussion forums. The conversation is nurtured and guided by the museum through the suggested themes, activities and tasks. Daycare educators are also encouraged to tell about their experiences, obstacles and suggestions for improvement.

Social media tools, as for example the services for sharing video, can be used in parallel and in this way ensure reliability and enhanced multi-platform exposure; for example the previously mentioned benefits of publishing content both to Vimeo, YouTube and ConnectedDay as simultaneous strategy to ensure the availability of our materials and to reach a broader audience.

In order to get the most out of the implementation of social media tools they need to be merged into the daycare and museum ecology. It has proved beneficial to plan and design the implementation of these social media tools in collaboration with the museum and the daycare personnel. The tools must be integrated with the existing know-how of museum docents and daycare educators. Content planning should also be a group effort. Through active collaboration these services could be a powerful means of promoting cultural education among younger audiences. They could also promote more opportunities for cooperation between daycares and museums.

In this paper we described how the needs of the project guided the choice of the services selected for sharing photos, documents and videos. Choosing the right social media services is part of the design process as they will lend their identity to the final project.

Our experiments with using video as a communication tool between daycares and the museum have led us to formulate the hypothesis of creating a museum Web TV channel directed at schools. There is inexpensive technology available to make this possible. Digital video is becoming an extremely popular medium. A wide range of photo cameras and most mobile phones already provide very reasonable video recording capabilities; and most laptop computers now come with an integrated webcam. In addition internet is quickly penetrating classrooms. This phenomena points out the need for further research on these issues.

The contents of a museum's Web TV channel are mostly video-on-demand that educators and children browse according to their needs. Video content focuses on presenting certain topics and proposing a series of activities and tasks, as in the case of *Oswald and the Objects*. Video contributions from the children and a set of social interaction tools, such as commenting, tagging and rating options and rapidly sharing links to other social networks are used to improve audience engagement and content distribution. Although we acknowledge that more work needs to be done for developing the concept of a Web TV channel for museums and daycares, our project explores some of the possibilities and obstacles that this type of channel could provide.

5. ACKNOWLEDGEMENTS

We thank the daycares and children who participated in the workshops. We are grateful to Ilpo Kari and all colleagues in Media Lab, Design Museum and Zipipop for their collaboration and assistance. *Oswald and the Objects* was supported by AVEK's Digidemo grant. Special thanks to Andrea Botero for encouraging us to write this paper, Hanna Kapanen for valuable comments and Richard von Kaufmann for final revisions.

6. REFERENCES

[1] Design Museum (2007). Design Museum website. Retrieved February 2, 2010, from http://www.designmuseum.fi/

[2] The Secret Life of Objects (2007). Blog of the project. Retrieved July 12, 2009, from http://thesecretlifeofobjects.blogspot.com/

[3] Oswald and the Objects (2009) Website of the project. Retrieved January 25, 2010, from http://www.esajaesineet.com/

[4] ConnectedDay (2009). Website of a daycare service provider. Retrieved January 9, 2010, from http://www.connectedday.com/

[5] Fisher, M., et al., *The Art of Storytelling: Enriching Art Museum Exhibits and Education through Visitor Narratives*, in J. Trant and D. Bearman (eds.). Museums and the Web 2008: Proceedings, Toronto: Archives & Museum Informatics. Published March 31, 2008. Consulted September 28, 2009.http://www.archimuse.com/mw2008/papers/fisher/fisher.html

[6] Salgado, M., Jauhiainen, T., and Díaz, L., 2009. *Re-thinking an Annotation Tool for Museum Community-Created Content*. Nordic Design Research Conference. Nordes, August 30- September 1, Oslo.

[7] Scribd (2010). Website for document sharing. Retrieved July 9, 2008, from http://www.scribd.com/

[8] SlideShare (2010). Website for document sharing. Retrieved January 21, 2010, from http://www.slideshare.net

[9] Flickr (2010). Website for photo sharing. Retrieved January 25, 2010 from http://www.flickr.com

[10] Vimeo (2010). Website for video sharing. Retrieved July 9, 2008, from http://vimeo.com/

[11] YouTube (2010). Website for video sharing. Retrieved July 9, 2008, from http://www.youtube.com/

[12] Design Squad (2008). Website of the project Design Squad. Retrieved January 30, 2010, from http://pbskids.org/designsquad/

[13] DragonflyTV (2008). Website of the project DragonFlyTV. Retrieved January 30, 2010, from http://pbskids.org/dragonflytv/

Tele Graph: Information Enriched Interactive WEBTV

Mehmet Özer Metin
AnelArge, Turkey
Hacettepe Üniversitesi Beytepe
Kampüsü, 2. Arge Binası No:11
+ 90 312 2992630
ozer.metin@anelarge.com

Siyamed Sınır
AnelArge, Turkey
Hacettepe Üniversitesi Beytepe
Kampüsü, 2. Arge Binası No:11
+ 90 312 2992630
siyamed.sinir@anelarge.com

ABSTRACT
With the emerge of digital convergence concept, conventional TV watching experience has been shifted from traditional TV's to WEBTV. However in the web environment, users claim new demands such as interactivity, information enrichment, and personalization. In this paper, an information enriched interactive WEBTV platform has been presented. The first phase of mentioned environment has been developed for a Turkish Telecom Company and went live under Tivibu (http://www.tivibu.com) brand. The main motivation for this work is building a new step up to traditional TV streaming with interactivity using community services such as Facebook, Twitter and enrich online TV stream with additional information from the semantic data stores using LinkedData.

Categories and Subject Descriptors
H.3.5 [**Information Storage and Retrieval**]: Online Information Services. D.2.6 [**Software Engineering**]: Programming Environments - Interactive environments. H.3.3 [**Information Search and Retrieval**]: Information filtering.

General Terms
Design, Experimentation,

Keywords
Internet TV, Semantic Web, Recommendation Engines, LinkedData, Personalization, Social Networks.

1. INTRODUCTION
TV watching experience has been shifted to web environment significantly over a few years. According to Nielsen study [1], the average consumer spent four hours on the Internet and 22 minutes watching online video. Internet video consumption is also on the rise, with users watching 53 more minutes of video online in Q3 2009 versus the same period in 2008.

Another statistic from the Nielsen research is that social networks are becoming a popular source of online video. Measured overall, the amount of time spent watching social network videos increased 98 percent from October 2008 to October 2009. Smaller increases were measured in other age groups. The 35-49 demographic increased its viewing of social media videos by 37 percent and the 65-plus age demographic increased viewing by 47 percent when compared to last year.

The study concluded that in 2008, Americans consumed information for about 1.3 trillion hours, an average of almost 12 hours per day. The study's authors say that amounts to 3.6 zettabytes and 10,845 trillion words. Measured on per-person basis, that amounts to 100,500 words and 34GB of data.

By the help of these statistics, one can conclude that traditional TV watchers tends to be more demanding about not only watching TV and video on Internet but also requires interactivity using social networks and more information enriched TV programs. We can state that traditional TV will be evolved into interactive personalized information and entertainment platform on Internet, covering 1) Interactive information services, 2) Interactive entertainment services, 3) Interactive communication services using social networks, and 4) Personalization of content and advertisement services.

Following this vision, we have proposed and implemented first phase of Tele-Graph: Information Enriched Interactive WEBTV for the Turkish Telecom Company which has been commercialized under the brand of Tivibu (http://www.tivibu.com). Currently, the system consists of over 80 national and international TV channels and over 300 Video on Demand content. The intended users are all ADSL subscribers which is a number over 5 billion.

Upon classical WEBTV services like, live TV and VoD streaming, transactional and subscription based VoD purchasing, program reminder and EPG filtering, Tele-Graph system builds an interactive entertainment platform that enables information enrichment services about current program. Interactivity has been achieved by the help of social community services like Facebook and Twitter where users can update their status according to program that are currently watching, share their opinions, recommend programs and see which programs that have been currently watching by their friends. This way, Tele-Graph enhanced classical TV watching experience to community-based experience using social community services. Information enrichment has been accomplished using EPG of a programme in TV-Anytime format and LinkedData. The EPG data is enriched from sources like IMDB, and then key phrase extraction is executed on this data. Having the key phrases of the programme, the system gets new information about those key phrases using LinkedData. Moreover, users can see IMDB information about any live program or VoD content. The user actions on the

displayed information are tracked and this tracking information is used to tune the key phrase extraction and LinkedData navigation.

The paper is organized as follows: section 2 will describe current Tele-Graph infrastructure, section 3 will concentrate on community service integration and information enrichment services, section 4 will described the future works for Tele-Graph project and concludes the paper.

2. TELE-GRAPH INFRASTRUCTURE

The overall system architecture is given in Figure 1. The system consists of 6 main components, namely; Streaming Infrastructure, Client Services, Semantic Service Server, Web Client, Content Management Servers and WEBTV Middleware.

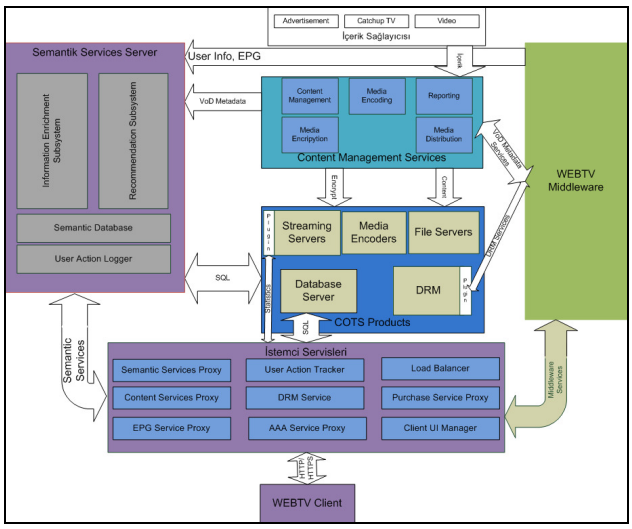


Figure 1 Tele-Graph Infrastructure

2.1 Streaming Infrastructure

Streaming Infrastructure is distributed to 6 different streaming centers located in 5 biggest cities of Turkey. Total number of 120 streaming servers, running IIS Media Services, has been deployed on these sites. The streams of live TV channels have been secured by URL Signing and for all VoD content, DRM is activated using PlayReady servers. The streams are encoded using VC-1 Advance profile and located in ASF container. All streams use smooth streaming technology that enables adaptive streaming of media by dynamically monitoring both local bandwidth and video rendering performance, Smooth Streaming optimizes playback of content by switching video quality in real-time [Microsoft]. Currently three levels of bit rate; 400 kbps, 700 kbps, 1 Mbps are available.

2.2 Client Services

Client Services have been designed as a service platform that serves necessary services for the clients. It can implement and serve its own services and also proxy remote services that have been served by middleware. The main services are EPG service proxy, DRM service proxy, load balancer service, AAA service proxy, purchase service proxy, content service proxy, semantic service proxy, preferences and user action tracker services.

2.3 Semantic Services Server

The main concept of this paper, information enrichment, is implemented in semantic service server. It contains recommendation subsystem, information enrichment subsystem and semantic database. The recommendation subsystem tries to match dynamically updated user profiles according to user implicit and explicit feedbacks to TV programs and VoD content. The recommendation subsystem details are out of this paper scope and information enrichment subsystem will be described in section 3.2.

2.4 Web Client

The client application of Tele-Graph project was implemented using Silverlight technology [Microsoft Silverlight] that enables rich internet applications running on browser environment. The web client contains a smooth streaming media player, necessary UI's for services like EPG, VoD, etc. and interactive elements like Facebook and Twitter integration that will be described in section 3.1.

The other elements that complete whole system are middleware that serves basically user management, subscription and purchase services; content management services that is responsible for media encoding and management of VoD content. Figure 2 shows a screenshot for Web Client UI while the user also connected with Facebook.

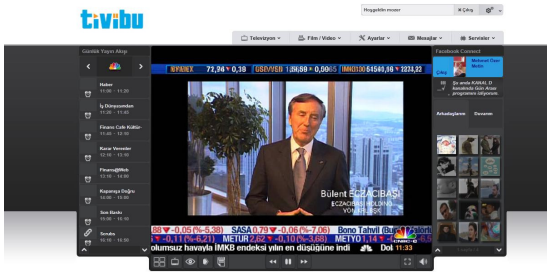


Figure 2 Web Client UI with Facebook Integration

3. INTERACTIVE AND INFORMATION ENRICHMENT SERVICES

3.1 Interactive Services

The main motivation behind interactivity of Tele-Graph project is not building a new and unique community under WebTV domain but build a sub domain regarding with TV watching experience using existing community services. Facebook integration is one of the earlier implementation of this idea. Facebook integration is mostly handled on client side where the WebTV client application connects to the Facebook using Facebook Connect API [2]. The communication is handled using Facebook JavaScript API, as well as cross domain communication channels. Facebook query language FQL and Stream objects were also used for retrieving user information, wall, friend list, etc.

There are four main user scenarios over Facebook integration;

(1) Users can update their status according to their current watching program. The system automatically fills the status with a message like **"**I am current watching "Star War II: Attack of the Clones" film on "Movie Channel"**"** and allow user to alter this message. The

user can also drag-drop any program on Facebook icon so that he can publish a story about selected program. The system also retrieves related program and channel info and automatically generate template status message.

(2) Users can see their friend status and as well as if they are also Tivibu users, they can see which channel and which program that their friends are currently watching. They can also see the watching history of their friends. This use case requires acceptance of the users whether they want their program watching history to be seen or not.

(3) Users can see their Facebook wall on Web client. They can retrieve any story, picture and video on their wall. Since most of the sharing in Facebook depends on video content. The player of client application can also start and show Facebook video streams.

(4) The users can became a Tivibu fan and make discussions about any program or video content using client application on Facebook.

Interactive services aims to change traditional TV watching experience to a community based experience using mostly used online community and social services. Facebook is one of these services where the client application can allow many of them such as Twitter, Digit, etc.

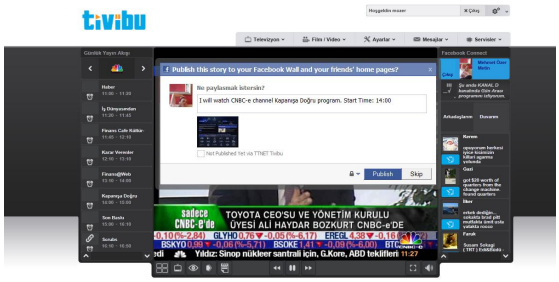


Figure 3 Publishing a story about program on Facebook

3.2 Information Enrichment Services

What current Internet had built on was to provide links between text documents which would improve navigation and learning. This way it was possible to provide a link to a concept, idea or a fact in another document; where a reader who would like to learn about that concept would easily jump to the document describing the concept. The same approach is valid for multimedia content such as videos. Even before Internet, Hyperland [3], a documentary made by Douglas Adams in 1990, mentioned that current TV systems lack of additional information and interactivity.

As Internet provides a huge amount of data about lots of concepts, now that it is possible to enrich TV content automatically with textual and multimedia content as Zemanta [7, 8] does for blogs. Tele-Graph applies what Zemanta does to blogs to videos streaming to live video. In Tele-Graph, while a program or movie is displayed, additional links, images, videos and descriptive hypertext about the concepts in the video is also displayed. Figure 4 displays the layout of TV stream and additional information on the content of the stream.

Advances in LinkedData [4, 5] enable realization of automatic runtime enrichment of TV content. LinkedData is the next step occurring in the Internet on the path to "Internet of Things" [6]. As an extension of Semantic Web, LinkedData provides methods of exposing, sharing and linking data on the Web using dereferenceable URIs. Currently there are many providers of LinkedData where DBpedia [9] (a provider that extract data from Wikipedia) and Freebase [10] are the biggest in size of concepts they are covering. The data provided by a LinkedData provider is linked with other data in the same provider, but also with the data of different LinkedData providers. There are open data about many concepts including movies, species, people, organizations, locations an etc. RDF [11] is used to describe the data and SPARQL [12] is used to query that data from the endpoints of data providers.

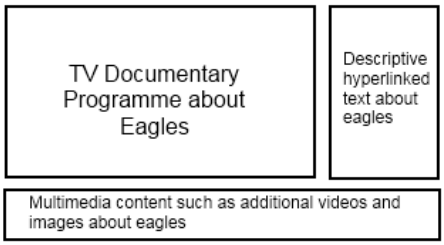
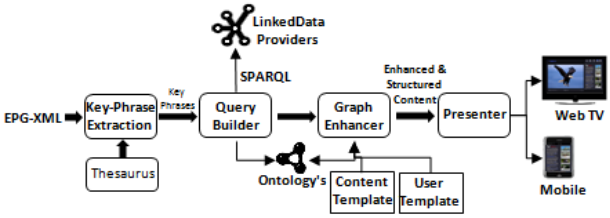


Figure 4 TV Layout for Information Enrichment

Tele-Graph is heavily built on LinkedData and the name is inspired from the "Tele" from television and "graph" from the graph of data present in LinkedData. In Tele-Graph, DBpedia is used as the primary data provider and links to some other concepts on different providers are also used (such as Linked Movie Database).

The method for information enrichment works as follows. First of all a Thesaurus is constructed in order to be able to train Key Phrase Extractor (KPE). Thesaurus is constructed using the concepts already exists in LinkedData providers. The data is queried against: Concepts, name/alternative names of concepts; the categories and labels of categories; the related concepts that fall into the area of focus and redirects (in Wikipedia terms). Using all of the labels and URIs, a thesaurus in SKOS [13] format is constructed. While constructing the SKOS thesaurus for each concept, the redirects and alternative labels are used as "altLabel" construct of SKOS. Each concept is linked with categories using "narrower" construct of SKOS. Each related concept about a concept is linked via "related" construct of SKOS.



In order to be able to create a learning model for KPE, sample learning documents where a set of text and the key phrases included in that text is prepared. Since the number of concepts in LinkedData is too big to construct training document-key phrase pairs manually, an automated method is used. In case of concepts that have been mentioned in DBpedia, the long abstracts and related Wikipedia articles are fetched as plain text. The key concepts in this plain text are assumed to be the concept itself

with all its related concepts and categories. Therefore labels of concept, related concepts and categories are provided as key phrases in Wikipedia article. Once the thesaurus and training text is prepared an enriched version of KEA [14] is trained against these training data. The reason to train KPE with LinkedData concepts is to output the LinkedData URIs as Key Phrases. Once the KPE is trained with thesaurus from LinkedData, it outputs the labels of concepts as Key Phrases which directly points some URI in the LinkedData cloud.

At execution time, for a program displayed on Web TV; the Electornic Programme Guide (EPG) data is fetched from EPG providers. EPG contains structured and well-defined data such as actors, program type and genre; whereas it also contains unstructured text descriptions about program. This text description is provided as an input to pre-trained KPE and the URIs to data in LinkedData providers are given as an output of KPE. Once nodes in the giant data graph is found some SPARQLs are executed against LinkedData providers which traverses the data graph and tries to find most related concepts with the current set of graphs (such as links between two nodes found in the first case). This step is named as "Graph Enhancer" which can be described as a semantic enhancement of found concepts. During the semantic enhancement process the process of finding most related concepts do change according to user profile and meaning of the edges between concepts.

Once graph enhancement is accomplished, the system ends with a set of nodes in graph and their attributes. These attributes and nodes are merged according to pre-defined templates for some ontology nodes in DBpedia such as Persons, Locations or Species. The final data includes images from different sites (such as Flickr), information from Wikipedia, albums in Amazon and home pages of concepts, as well as descriptions of concepts. Finally the structured data is provided to Presented which displays the multimedia information in WebTV or a 2nd screen such as a mobile phone [15].

4. CONCLUSION

In this paper, an information enriched interactive WEBTV platform has been presented. The first phase of this environment has been developed for Turkish Telecom Company and went live under Tivibu (http://www.tivibu.com) brand. The system consists of over 80 national and international TV channels and over 300 Video on Demand content. The intended users are all ADSL subscribers which is a number over 5 billion. The whole project is named as Tele-Graph where the name is inspired from the "graph" of data provided by LinkedData.

Although the whole project consist of many parts related with classical web TV application like; streaming, content management and delivery system, middleware, etc. In this paper only interactive services and information enrichment services has been focused.

The interactivity of the system mainly depends on Facebook integration where the users can interact with the system and their friends. This interaction is focused on forming community that share information about TV and video domain.

Tele- graph: information enrichment Web TV project also aims serving rich content that has been gathered using various Internet sources like LinkedData. The information enrichment is done automatically generating key phrases from EPG data and video metadata and finding relevant data that is build on these key phrases. The other part of information enrichment is presenting the content in an interactive and user friendly way.

5. ACKNOWLEDGMENTS

Our thanks to Hakan Çağlar and Fatih Orhan for their innovative ideas and supports.

6. REFERENCES

[1] Nielsen A2/M2: Three Screen Report, Q3 2009, http://en-us.nielsen.com/main/insights/nielsen_a2m2_three, Last Accessed 05/02/2010.

[2] Facebook Connect, http://developers.facebook.com/connect.php, Last Accessed 05/02/2010.

[3] Adams D., Hyperland, Documentary by Douglas Adams, 1990, http://video.google.com/videoplay?docid=7190175107515525470&ei=qVFpS_X2OJuM2AKY2qC2CA&q=Hyperland#, Last Accessed 05/02/2010.

[4] Bizer C., Heath T., Berners-Lee T., Linked Data - The Story So Far, International Journal on Semantic Web and Information Systems, Special Issue on Linked Data, 2009.

[5] Bizer C., Heath T., Ayers D., Raimond Y., ESWC2007 Poster Paper on Linking Open Data, 2009.

[6] Commision of European Communities; Early Challenges regarding the "Internet of Things", 29.09.2008.

[7] Mihaila C., Zemanta: A Content Recommendation Engine, http://www.slideshare.net/claudiumihaila/zemanta-a-content-recommendation-engine, Last Accessed 05/02/2010.

[8] Zemanta Web Site, http://www.zemanta.com, Last Accessed 05/02/2010.

[9] Jentzsch A., DBpedia – Extracting structured data from Wikipedia. Presentation at Semantic Web In Bibliotheken (SWIB2009), 2009.

[10] Freebase, http://www.freebase.com

[11] RDF, W3C Standard for resource description, http://www.w3.org/RDF

[12] SPARQL, W3C Standard query language for RDF, http://www.w3.org/TR/rdf-sparql-query

[13] SKOS, Simple Knowledge Organization System, W3C Standard to support the use of knowledge organization systems (KOS) such as thesauri, classification schemes, subject heading lists and taxonomies within the framework of the Semantic Web, http://www.w3.org/2004/02/skos

[14] Medelyan, O., Witten I. H. "Thesaurus Based Automatic Keyphrase Indexing."In Proc. of the Joint Conference on Digital Libraries, Chapel Hill, NC, USA, pp. 296-297, 2006.

[15] 2 Screen solution, form of interactive TV that enables information about a TV show to be accessed via the internet on a mobile phone, laptop or desktop PC; http://en.wikipedia.org/wiki/Two-screen_solutions

Interactive TV-centric Health Care Services in Smart Homes

Pau Arce
Universidad Politécnica de Valencia
paarvi@iteam.upv.es

Francisco Fraile
Universidad Politécnica de Valencia
Interactive TV Arena
ffraile@iteam.upv.es

Juan Carlos Guerri
Universidad Politécnica de Valencia
jcguerri@dcom.upv.es

ABSTRACT
This paper approaches TV-centric health care services provisioning in Smart Home environments. First, users perform exercises and other activities monitored through unintrusive sensors (surface electromyography, positional sensor). The tests also include questionnaires in order to allow a therapist to assess the health condition and the habits of the patients at home. The evaluation of the therapist is used to generate media recommendations with clinical advices and new exercises that continue with therapy.

Categories and Subject Descriptors

H.5.2 [User Interfaces]: User-centered design, Evaluation and Methodology, H.3.4 [Systems and Software] Distributed Systems, D.2.11 [Software Architectures]: Domain-specific architectures,

General Terms
Design, Experimentation, Human Factors, Verification.

Keywords
e-Health, IPTV services, usability.

1. INTRODUCTION
Healthcare practice benefits from the latest advancements in Smart Homes related technologies. Ambient sensors in the home collect information about the environmental conditions wherein patients live. Wearable biometric sensors in furniture and clothes connect to the home sensor network to monitor the patients' health conditions. Accelerometers, electronic weight scales or pedometers contribute to the plethora of information useful to monitor the health habits of patients at home. Thanks to secure broadband connections, all this information can be stored in remote databases for latter processing and analysis.

Furthermore, communication services bridge the gap between therapists and confined patients. Social networks ease the sense of alienation, allowing patients to build communities and share experiences. Lastly, modern internet TV and interactive TV technologies enable the offering of these services in a friendly manner, with the TV as the central User Interface, interacting with other multimedia devices in the Home Network.

With these technologies it is possible to provide remote assistance to patients at home in an effective and unintrusive manner. As a result, the costs of providing remote health assistance to confined citizens are drastically reduced. Moreover, the tracking of user activities at home allows specialists to assess the health habits of patients very accurately, thus improving the results of the treatment. The cut back on costs, together with the enhanced functionality, motivates the development of general-purpose applications that do not only aim to treat specific diseases but that also try to encourage healthy habits in society. In the context of massive user communities, it is economically viable to produce health related media content associated to the application, in order to create an interactive TV concept that integrates all this different application components: home sensors, Electronic Health Record (EHR) databases, interactive TV and personalized media.

This paper presents a telemedicine application that collects health records from users at home and stores the data in an EHR database for further analysis. A medical specialist evaluates the result of the activities and provides a diagnosis. The system implements a recommender engine that presents to patients at home information related to their treatment: videos with healthy habits, new exercises or questionnaires. The application also implements video communications between patients and therapist.

The rest of the paper is structured as follows: next section presents the system architecture supporting the service. Section 3 focuses on the application and the user interface. Section 4 presents results and conclusions.

2. ARCHITECTURE
2.1 System Design
The system platform allows therapists to obtain health records from patients at their homes, avoiding visits from patients to a medical center or visits from therapists to patients' homes. Moreover, the medical reports, diagnosis and subsequent treatment steps are available on patients' interactive TV sets.

Figure 1 depicts a diagram of the system under development. The figure shows three differentiated domains: the Smart Home domain, the service provider domain and the therapist domain. The following list summarizes the steps involved in the functioning of the application:

1. The therapist draws the therapy work plan, which will be stored in the server database. Each work plan consists of questionnaires and multimedia content describing medical advices and therapy exercises. The questionnaires are used to create and update user profiles.
2. The patient receives the work plan on his TV set as an interactive TV service. It is necessary that users login to the service and fill in the initial questionnaire on their interactive TV set in order to access their personalized work plan.
3. When needed, the sensor equipment undergoes a remote configuration to prepare for the activities. Also, users receive (video) configuration instructions for the positioning of sensors or the execution of exercises.

4. Once the patients are ready, they perform the exercises and the data (biometric signals, weight, blood pressure) capture starts. The application sends the data back to the server. The data can be captured and sent to the server in real time.

5. The server processes the raw EHR data to detect anomalies and assess a preliminary diagnosis. Together with the questionnaire, the EHR data is used to create user profiles that describe the state of patients and their health habits.

6. Therapists access the data in the server to follow up the progress of each patient. The server user interface presents therapists a comprehensive summary of the results of the tests and questionnaires, the preliminary diagnosis and a list of media items related to the patient profile. Then, the therapists generate the pertinent medical reports.

7. Again, patients receive the results of the tests, together with the medical reports, the updated work plan and the media recommendations on their interactive TV sets.

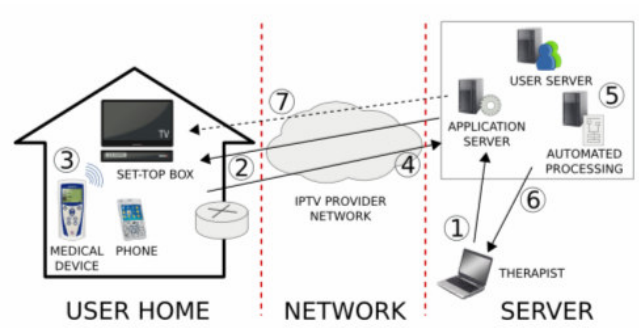


Figure 1. System architecture

The service requires a broadband connection to the home and a TV device able to render multimedia content from the broadband media (i.e. a Hybrid Broadband/Broadcast Receiver). If the receiver is an IPTV Set-Top Box (STB) connected to the Home Network, then the service requires of an adequate IPTV System configuration in order to enable access to the service from STBs.

2.2 Media Recommendations

As mentioned before, the server produces media recommendations. The purpose of the media in the service is not only to depict exercises related to the therapy, but also to promote healthy habits acquisition among patients. The acquisition of good daily habits has a great impact on people´s health condition. Moreover, unhealthy daily habits normally cause or aggravate health problems. Therefore, it is important to also care for the behavior of patients, correcting unhealthy habits and encouraging the acquisition of healthy habits.

For this reason, the system profiles the patients' daily habits as well as physical condition. To create the patient profile, the system can use implicit information, obtained through questionnaires, or explicit information, obtained through sensor data.

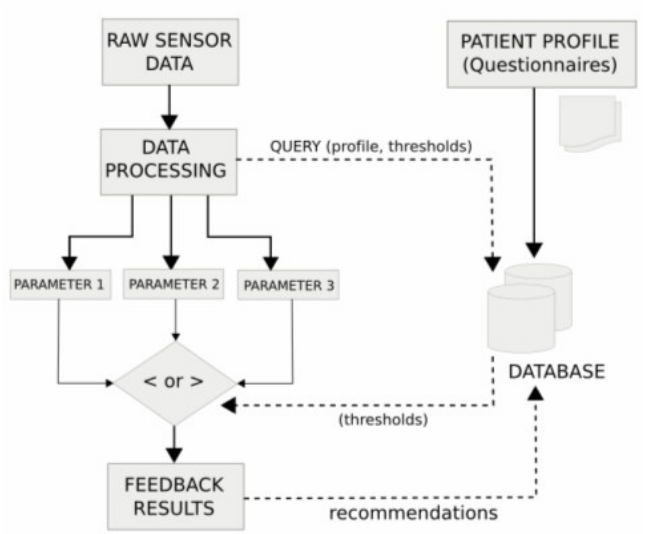


Figure 2. Diagram of the recommender

The sensor data is processed in order to detect pathologies in patients (e.g. a blood pressure measurement can be used to detect unhealthy blood pressure levels) or patterns on the patients' habits (e.g. a position sensor can be used to detect low mobility of confined patients). Sensor data processing normally implies comparing measurement levels with certain thresholds as shown on Figure 2. The questionnaires are used to determine how healthy the habits of patients are, as explained in section 4.1.

The user profile is described in a domain specific ontology that represents the different health conditions and the patients habits related to their treatment. The multimedia content is described with the same ontology, defining the target profile of patients and the healthy habits that they promote. Recommendation algorithms can then determine which videos are better suited for a specific patient.

3. APPLICATION PROTOTYPE

The service has been prototyped and tested in laboratory conditions[1]. During the user trials, the use case under study was the treatment of low back injuries and their relationship to unhealthy habits. Low back disorders are a major health problem in western countries. It is estimated that 58-84% of the population will suffer back pain at some point of their lives. Recent research works evidence that the lumbar vertebrae undergo significant stress when persons prolong uncomfortable postures, either standing, sit or laid, during long periods of time. Other major reasons are bad positions kept when lifting or transporting weight. It is clear that the acquisition of healthy daily habits can reduce drastically the amount of back injuries. The service prototype aims at collecting data to evidence possible injuries or diseases that cause the pain and at identifying daily habits that may cause or deteriorate the patients´ condition.

[1] Telefonica and the Spanish Ministerio de Educación y Ciencia within the MIQUEL (TEC2007-68119-C02-01/TCM) project sponsored the development of the prototype.

The therapist uses the service to detect low back injuries, through electromyography data and to detect unhealthy habits that may cause the injury, through questionnaires and through exercises monitored with electromyography. The server identifies anomalies in the results of the tests. Together with the results of the questionnaires, the electromyography results are used to create a profile of the patient low back condition and related daily habits. A media recommender finds media content useful to reinforce the patients´ awareness of healthy habits related to low back pain and related injuries, encouraging healthy habit acquisition among patients. Therapists supervise the results and generate reports.

A number of applications have been developed in order to provide the platform with this functionality. On one hand, an interactive TV application implements the interface between the platform and patients at home. Through the TV set, the patients can visualize the work plan, fill in the questionnaires and review the medical reports and the media recommendations.

On the other hand, an application installed on the mobile phone is in charge of communicating with the medical devices in range, collecting data and then, sending the gathered information to the server. The mobile application is configured to operate with the medical device used on the prototype but is easily extensible to other wireless medical devices with open specifications.

All medical data picked up by the phone application can be sent immediately to the server in a real-time manner or they can be stored for further transmission if no connection is available at that time.

On the server side, the Application Server implements user interaction with the system through the IPTV platform. Questionnaires, exercise videos, measurement information and recommendations, as well as medical reports, are stored and served by the Application Server, which offers Web Services so mobile phones or medical devices in the patients home can send the measurement data. Moreover, user authentication is carried out by the User Server, where user profiles and medical data are safely stored. In addition to these components, the automated Processing Server provides the recommendation engine. Also, it is responsible of conform a smart response from the medical data, answered questionnaires and user profiles.

The therapist access to the system to examine patients' data, write medical reports, design new questionnaires and recommendations or just review the automatically generated recommendations.

The equipment used during the trials is listed below:

- Set-Top Box: A Hybrid IPTV/DVB-T STB with an embedded browser and Ethernet connection.
- Medical Devices: The medical device is equipped with surface electromyography sensors, which are very easy to apply. Patients only need to follow the video guides available on their TV sets. This device is also equipped with a wireless interface in order to transmit the measured data to the phone in real-time.
- Mobile phone: When needed, mobile phones are used as gateways that communicate with medical devices in order to capture health data and send it to the server.

3.1 User Interface

The user interface has especially been designed for the TV platform, adapting graphical components' size to the common TV screen resolutions. Furthermore, user interaction is carried out through a remote control and some operations could become tedious and confusing. For this reason, user interface has been designed taking the TV media into account and some of the adopted solutions are limiting the number of selectable items, incorporating an on-screen keyboard and allowing fast navigation through the color buttons.

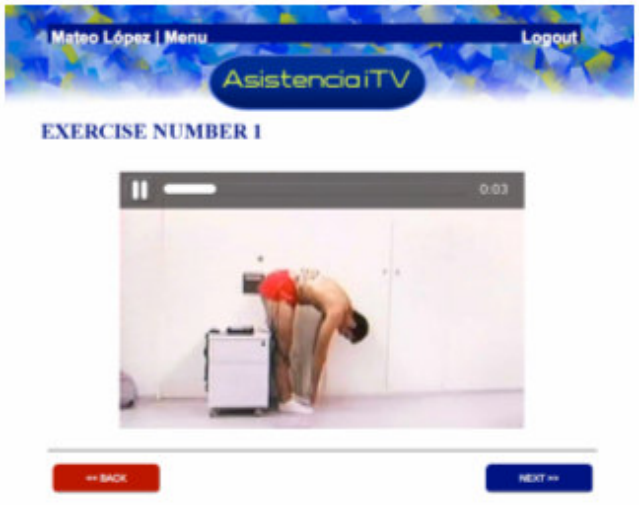


Figure 3. Screenshot of the application prototype

4. RESULTS

The first results generated evaluate how reliable the initial questionnaire is when assessing whether the users' habits at home are healthy or not. For every application domain, it is necessary to produce a questionnaire that can discriminate patients with healthy habits from patients with unhealthy habits, as well as which habits should be changed. Section 4.1 explains in more detail the problematic of producing good questionnaires for the assessment of habits related to low back disorders, together with the results of the evaluation of the questionnaire.

The usability of the TV UI (User Interface) has also been tested, by means of end user trials conducted in laboratory conditions with the application prototype. It is clear that the success of these applications greatly depends on their usability, feasibility, and performance of transport services offered by the communication networks. Also, the characteristics of TV interfaces are very different from traditional PC web browsing. In the literature there are several works about requirements for designing handheld computers systems [4], ubiquitous computing evaluation [5] and more specifically, focused on experimental performance evaluation of communication networks offering health services [2], [7].

4.1 Tests and Questionnaires Reliability

The methodology used in back ache programs for health promotion has been questioned [3] and the results provide irrelevant information [1]. In order to address the problem of habit acquisition appropriately, it is necessary to assess accurately all

habits related to back health and to carefully establish a work plan to encourage the acquisition of healthier habits.

In order to assess the habits related to back health, we developed a new Likert questionnaire with 31 items and 5 levels of agreement. The questionnaire evaluates patients' habits in six different situations. There are four questions about the position of feet and the usage of footwear. Nine questions about the habits of sitting for reading, studying or in front of a computer. There are 12 questions about the carriage of weight and six about rest habits. The questionnaire was handled twice in a 4 week interval to a total of 100 persons. The total useful data came from 77 persons (38 man and 39 women). In order to analyze the validity of the questionnaire, the Test-Retest method and the Spearman correlation coefficient were used.

The first questionnaire obtained a mean score of 2.91 (DT=0.41). The second provided a mean score of 2.82 (DT=0.43). The Intra-Class Correlation Coefficient (ICC) obtained was 0.78, showing an acceptable stability. The following table (Table 1) shows the Rho Spearman values between the different the two instances of the questionnaire for every habit under evaluation:

Table 1. Spearman Correlation Coefficient between the two instances of the questionnaire

Habit under study	Rho Spearman
Standing	(0,36 - 0,56)**
Sitting (Leisure)	(0,34 - 0,53)**
Sitting (in front of PC)	(0,36 - 0,54)**
Carrying backpack	(0,30 - 0,54)**
Carrying weights	(0,36 - 0,46)**
Laying down	(0,62 - 0,85)**

****p< 0,001.**

The Spearman coefficients indicate a weak to moderate correlation in most of the habits under study. The three with the lowest correlation values were discarded from the questionnaire. The results show that the questionnaire is reliable as a global indicator to discriminate patients with healthy habits from patients with unhealthy habits. However, the low correlation coefficients indicate that the questionnaire does not show clearly the problematic habits and it is necessary to continue with further questionnaires to get an accurate user profile.

4.2 Usability Assessment

Usability data were collected using questionnaires at the end of the trial. The tests were completed by 7 experienced users in communications technology (with knowledge about the usability concept). The questionnaire was based on [6] and [4], and included such items as attention, trust, conceptual model, interaction, invisibility and impact. The answers were either yes/no or chosen from a Likert scale, ranging from 1 to 5 (lower is better) (Table 2). The results showed a good satisfaction about the usability of the application, especially with the functionality and interaction.

Table 2. Usability evaluation of the application. Likert scale: from 1 (Very Correct) to 5 (Very Incorrect).

Variable	Mean	Sd
Attention		
Time spent using the application interface (m)	12	1.8
Trust		
Feel info was private (1–5)	2.28	0.7
Conceptual model		
Function as expected (1–5)	1.28	0.5
Interaction		
Disrupting (1–5)	1.85	0.4
Frustrating (1–5, higher is better)	4.57	0.5
Invisibility		
Understand key functions (1–5)	1.71	0.5
Guidance using the application (1-5)	2	0.8
Impact		
Motivating (1–5)	2	0.8

5. ACKNOWLEDGMENTS

This work was supported by the Spanish Ministerio de Educación y Ciencia within the MIQUEL (TEC2007-68119-C02-01/TCM) project and Cátedra Telefónica de la UPV. The company Interactive TV Arena provided the IPTV equipment.

6. REFERENCES

[1] Brattberg, G. 2004. Do Pain Problems in Young School Children Persist into Early Adulthood? A 13-Year Follow-up. In European Journal of Pain 8(3): 187-199.

[2] Chu Y. and Ganz, A. 2004. A Mobile Teletrauma System Using 3G Networks, in IEEE Transactions on Information Technology in Biomedicine, vol. 8, no. 4, pp. 456-462, December 2004.

[3] Nentwig, C. G. 1999. Effectiveness of the back school. A review of the results of evidence-based evaluation. In Orthopade, 28, 958–65.

[4] Palmblad, M. and Tiplady, B. 2004. Electronic diaries and questionnaires: designing user interfaces that are easy for all patients to use. Quality of Life Research, pp. 1199–1207, Kluwer Academic Publishers.

[5] Scholtz, J. and Consolvo, S. 2004. Towards a discipline for evaluating ubiquitous computing applications. Intel Corporation.

[6] Tsai, C. C., Lee, G., Raab, F., Norman, G. J., Sohn, T., Griswold, W. G. and K. Patrick. 2007. Usability and Feasibility of PmEB: A Mobile Phone Application for Monitoring Real Time Caloric Balance. Mobile Networks and Applications, vol. 12 , issue 2, pp. 173 – 184.

[7] Viruete Navarro, E. A., Ruiz Mas, J., Fernández Navajas J. and Peña Alcega, C. 2006. Performance of a 3G-Based Mobile Telemedicine System, in Consumer Communications and Networking Conference, vol. 2, pp. 1023- 1027.

Derrick Metadata Platform
An Internet Platform for Television Metadata

Alberto Attilio Reggiori
Derrick Media Ltd
3rd Floor, East Unit
12 Bridewell Place
London EC4V 6AP United Kingdom
alberto@derrickmedia.com

ABSTRACT
This paper provides an overview of the Derrick Metadata Platform, an Internet platform for television metadata under development at Derrick Media. The system aims to provide a white-label and open television metadata Software-as-a-Service (SaaS) platform for broadcasters, content owners and advertisers making metadata provisioning more audience ready leveraging the Internet and its model for collaborative knowledge creation. The aggregation, management and syndication of high quality media metadata is vital to increase revenues, easy repackaging and multi-platform delivery, licensing content and provide a better view of its core information assets. The paper major contributions include: (1) Introduce more accountability for metadata at the level of other entertainment assets (2) Promote openness and standardization of metadata (3) Create a controlling and reporting system around metadata to increase quality (4) Connect enterprise systems that produce metadata into one aggregation system that drive publishing on every output platform.

Categories and Subject Descriptors
H.3.5 [INFORMATION STORAGE AND RETRIEVAL]: Online Information Services.

General Terms
Design, Management.

Keywords
Enterprise Metadata Management, Information Integration, Television Metadata, SaaS, Workflow Management, Quality Management, Aggregation, Syndication.

1. INTRODUCTION
The world of television has been undergoing a major transformation over the past years from mass take-up of time-shifted viewing models, to on-demand programming, to the availability of video content on non-TV platforms and new commercialization opportunities such as Internet advertising. At the same time an everyday exponentially growing amount of programming is made available to the consumers via the Internet, setting major challenges for the television industry and undermining its business models. Broadcasters, content owners and advertisers must use every advantage at their disposal to reach underserved audiences.

Programme metadata, descriptive information about television programmes, is the key enabler to monetize content and create value, helping search, discovery and navigation of products, creation of higher user engagement experiences, accurate reporting and analytics. Metadata is an important contributor to saving time, costs and effort where productions have to re-author and output material to more distribution channels, leading to better services and more satisfied customers. Ultimately metadata impacts the entire media content value chain and it is as important as the actual programme itself, or put it another way, *metadata is a product itself*.

Meanwhile technologies known collectively as Web 2.0 have spread widely among consumers over the past five years. Social-networking Web sites, such as Facebook [1] and Twitter [2], now attract more than 100 million visitors a month. As the popularity of Web 2.0 has grown, television industry has noted the intense consumer engagement and creativity surrounding these technologies. In this new model, consumers, or the audience, have a way to contribute information to a community system about a programme, making it possible to change the way metadata is created and consumed.

Media organisations, on the other hand, have to integrate various legacy software systems in order to retrieve, and consolidate information in house. Today media organisations have very little, if any reports, about the usage of metadata downstream in programme delivery. This is an information vacuum that affects the interaction between broadcasters, producers, and advertisers.

Metadata, when made available in an open and accessible form will not only help content creators understand how their content is consumed, but also provide a feedback mechanism that over time can improve accountability, quality and attractiveness, audience targeting, and promotion for programmes.

The Derrick Metadata Platform aims to provide a white-label and open television metadata platform for broadcasters, content owners and advertisers making metadata provisioning more audience ready leveraging the Internet and its model for collaborative knowledge creation.

2. CHALLENGES
Inside the media life cycle various fragments of television programme metadata are collected, processed, presented and disseminated by both TV platforms and various information systems. Such information is often hidden behind legacy and enterprise interfaces, stored in a not machine-processable form, and by nature fragmented in time and space. In fact, if one would take an hypothetical snapshot of metadata about a programme, it would be evident that information is incomplete and not self-consistent. There is not a time 0 from when a metadata record starts growing up and a correspondent time n when that metadata unit can be said to be completed. Information comes and goes

without any starting or ending point, and it is often available before and/or after the content they refer to is made available, not being tightly coupled with it.

It is essential then to outline the flows and define the processes that pertain metadata creation, manipulation, delivery and consumption, without paying too much attention to the actual metadata static consistency as a whole. Due to the multiplicity of stakeholders and actors participating in the metadata value chain, it is often difficult if not impossible, to devise a unique all-comprehensive model of programme information at given time; but instead different semi-structured versions of the same model (or views) must coexist[1]. This means that in order to deliver accurate, consistent and timely information to the right stakeholders at the right time, a process-driven approach to information sharing and application integration needs to be combined with a pure data-centric one.

In addition, most of the collected information about programmes often is missing "context" about for example where it came from, who created it, why etc. Likewise during programme metadata aggregation and management another key issue is the version control and conflict management of different versions of a metadata record; for example, when down the distribution chain one might not be able to distinguish where and why an error was introduced, or by whom. More generally, asynchronous collaboration and evolution of programme metadata can happen in several ways in the media life-cycle; with the risk that stakeholders cannot understand what has changed, and collaboration can quickly spiral out of control. Missed changes made by one stakeholder can unintentionally damage the work of others or even the entire metadata information record.

Thus for a broadcaster, content owner or advertiser it is essential to manage the flows and processes that pertain to metadata creation, manipulation, delivery and consumption in a way that is:

- completely transparent and natural
- does not pay too much attention to metadata static global consistency
- every stakeholder gets only and exactly what it expects to get
- preserve context of information
- support version control and conflict management

One important aspect is that media metadata legacy systems and infrastructures commonly have an archaic origin and a long history of improvements and extensions; and cannot generally be replaced from one day to another. No one wants to adopt new standards or protocols to take advantage of metadata augmented value systems.

A metadata platform really hits the mark when it is completely, seamlessly and painlessly integrated with content owners, broadcasters and advertisers their media life cycle business related processes, impacting their day to day work to the bare minimum; ideally not even think about metadata itself.

Furthermore during the production process metadata is often only informally collected by most media organisations, and to write good metadata is hard and costs money. This means that there is potential to build an efficient joined up lightweight metadata aggregation, management and syndication process that has the biggest potential for cost savings; where each metadata reuse point (E.g. avoid data re-entry and probably reduced errors) immediately turns in time saving, costs cutting and reduce people effort.

Ultimately getting programme metadata organized is a major problem for most broadcasters, content owners and advertisers. This is because metadata is managed by many groups (we call it the "IT problem"); or because there are too many similar systems not having a unified publishing platform. Or because various metadata information systems are missing an open access and standard interface; or they are even offline and not connected to the main distribution pipeline.

3. METADATA AS A PRODUCT

Programme metadata is descriptive information about television programmes, or data about data. Inside the media value chain, metadata varies in quantity, quality, granularity, descriptiveness, searchability and availability. Different stakeholders need to know different things about a specific product at different times; and metadata is used in different ways by different people depending on the media process their are involved into. Each workflow stage has it own specialistic metadata [3].

Most important, metadata is meaningful information in its aggregate - a single piece of metadata is merely a piece of detail data and in isolation is not very informative. Several pieces of metadata grouped and interrelated together are necessary to convey useful information objects about television programmes. Increasingly complex structures and substructures each with their own rules and encodings needs to captured and managed all the way through the metadata workflow.

The early capture of metadata provides downstream benefits providing context to content, and content consumption can become intelligent and interactive, providing more opportunities to engage audiences.

Metadata enables flexible ad logic, target advertising and new social media advertising opportunities by connecting people using social tagging, vocabularies and conglomerate those into communities having similar demographics. Metadata also enables television producers to provide data about their programmes in order to make them searchable by machines [4], and discover new related content; as well as drive better and more granular search results and new ways to navigate content. Using metadata broadcasters, content owners and advertisers can better understand how content is being consumed, and provides the feedback mechanism to iterate the cycle, improving targeting and messaging.

To accelerate metadata integration capabilities, media organisations need an integrated view of metadata and reports about its usage across entire enterprises including suppliers and audience. Metadata management, combining business process, business objects and operational information, would deliver the ability to improve the usability of metadata.

4. QUALITY OF METADATA

Due to metadata's pervasiveness, its quality is an important dimension for achieving the goal of excellent audience service. As the market for media grows, metadata helps provide broader,

[1] Another way to see this is the paradigm of shared vs. specialized metadata vocabularies: sometimes standardized meta- data fields are essential to a particular publishing vertical. But individual businesses, content, and systems also have their own needs.

richer source information so that the user experience is heightened and the media is easier to find, play and share.

Broadcasters, content owners and advertisers need to create a clear flow of metadata with quality control, audit trail, versioning and reporting features. This will hoist the personal and professional accountability for media metadata at the same level of quality as the media programme itself. Better metadata quality and quantity will bring improved navigation for consumer, speed and accuracy; as well as improved semantic definition and recognition to increase viewer share and hence revenues. This will also imply defining and using a set of standards for writing metadata, collecting metadata and using metadata on output platforms.

5. USER-CENTRIC METADATA

In the past decade a fundamental evolution in the attitude of consumers towards networked media has happened. The content infrastructure has changed from a static global library to a dynamic meeting place where users not only consume but share rich media content. Person to person communication mechanisms are reshaping forever the way people consume television content at home and they way interact with each other. Consumers actively contribute explicit data such as information about themselves, their friends, about the items they purchased, reviews, ratings, recommendations and comments about television programmes and related information.

Today, the online world has shifted to a model of collaboration and explicit data creation [5]. Inside social-networking Web sites, such as Facebook and Twitter, an immense web of human relations is taking shape. Millions of users are playing, working, flirting, and socializing online — and producing oceans of data. This has also affected the way media metadata is being created, managed, distributed and consumed inside the media value chain. While consumers are freely submitting their own information onto the web, the organizations that can effectively collect and analyze it and use it to build positive and successful relationships with consumers. Metadata is the key connecting tissue between broadcasters, content owners, advertisers and the consumer.

6. DERRICK METADATA PLATFORM

The Derrick Metadata Platform, here referred to as Derrick, is a white-label Software-as-a-Service (SaaS) platform for television metadata aggregation, management, visualization, analytics and syndication. By combining the best of hosted services, Open Standards and Social Media, it offers a completely integrated, secured and customisable metadata platform for broadcasters, content owners and advertisers. Derrick provides Open Application Programming Interfaces (APIs) to access, manage, distribute and consume television programme information through the whole pre/post-production, distribution and consumption processes.

Derrick enables media organisations to keep full control of their metadata assets, while consumers can better engage and interact with their rich media products. We provide full integration with major social media platforms such as Facebook and Twitter, allowing users to actively contribute information into Derrick about themselves, their friends, as well as reviews, ratings, tags, polls, favorites, recommendations and comments about television programmes.

Built bottom-up as a lightweight overlay to existing media organisations infrastructures, Derrick acts as a catalyst of media flows. Its internal fully extensible data model enables doing the best with existing metadata and structures, rather than superimposing an omni-comprehensive top-down data model through the whole metadata life cycle. Derrick's pluggable architecture allows to add specialized transformers and adapter modules for existing media formats, metadata standards, protocols and interfaces. Internally stored metadata is being fully contextualized, and an asynchronous change awareness framework is used for version control and dealing with conflict management.

In addition, a set of powerful applications and hosted services can be used by customers to manage quality of information, automatic validation and correction of errors, build reports and analytics about metadata usage; as well as create, edit and visualize metadata records using proprietary and standard metadata formats.

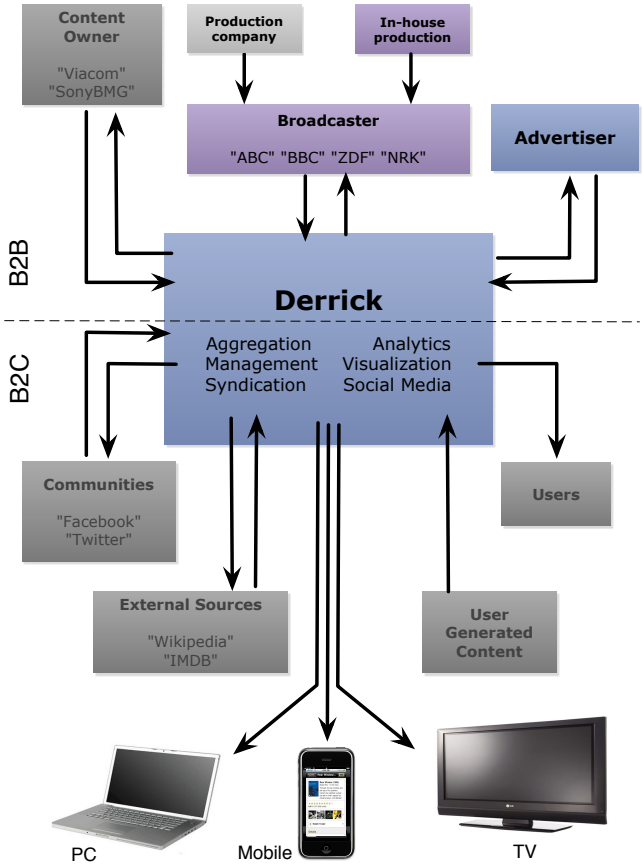


Figure 1 - Derrick Metadata Platform overview.

7. REFERENCE ARCHITECTURE

Derrick provides the following services:

- *Aggregation* - acquisition, collection, creation and sourcing of metadata
- *Analytics* - reporting and understanding of information
- *Management* - quality management, authoring and editing of rich metadata
- *Visualization* - display and visualize metadata structures and relationships
- *Syndication* - distribute metadata to partners, suppliers and consumers

- *Social Media* - user generated ratings, tagging, polls, quizzes, comments and activity streams

These services will seamlessly work together, adapting to media specific business workflows, helping to increase revenues, and scaling as metadata volume and traffic grow. In addition, they will be open for integration with best-of-breed technologies, including existing media metadata legacy systems and infrastructures or third-party applications. Figure 1 provides an high level overview of the metadata services provided.

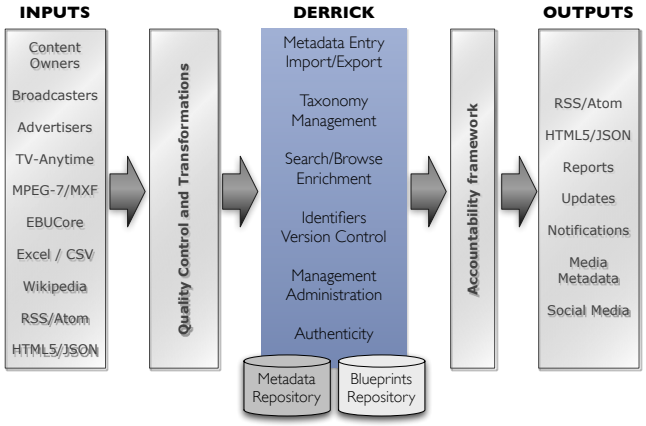


Figure 2 - High level architectural view.

Derrick offers a hosted enterprise-class environment. Internally the platform consists of a core set of components:

- *Metadata Repository* - a multi-tenant repository containing a consolidated and integrated view of metadata.
- *Blueprints Repository* - a repository containing business process blueprints (definitions).
- *Metadata Entry* - metadata collection management workflows, including insert, update, delete functions, and validation, versioning and duplicates detection.
- *Metadata Import/Export* - import/export of metadata created in approved metadata standards. Standards that are not supported can transformed or mapped accordingly by the user.
- *Taxonomy Management* - storage, maintenance and application of taxonomies across all the other components.
- *Search/Browse* - support for any combination of fielded metadata searching.
- *Identifiers Management* - persistent and unique identifiers are maintained for each media metadata resource being described.
- *Metadata Enrichment* - in order to provide a better view and navigation of programmes and resources, collected metadata needs to be augmented and enhanced.
- *Metadata Reporting* - in the drive to better accountability and in order to develop the reporting of metadata, which will connect the information with the responsible managers and contributors, a number of report types can be generated.
- *Management and Administration* - management and configuration of metadata fields, management of processes and workflows, management of user, group profiles and access levels as well as creation of programming interfaces to and reports from the platform, through these functions. In addition, these functions provides support of an audit trail over the lifecycle of each metadata record.
- *Authenticity* - interface with a rights management and/or access control system to allow or deny access to the stored metadata in the repository depending on access conditions specific to an object or set of related objects, user class, location etc.

In Figure 2 we provide an high level architectural view of Derrick.

8. APPLICATION SCENARIOS

Using the Derrick core set of metadata services and APIs it will be possible to leverage the Internet platform to program services to serve next generation television Web applications and services. Here is just a short list of example application ideas which could be developed on top of Derrick.

- *Next generation Programme Guide* - Application using the backend metadata platform services to find what's on TV, DVD and in theaters, and explore content by channel, director, cameramen or decompose metadata by time and place.
- *Programmes Support Website* - A semi-automatically built Website which a content owner or broadcaster could offer to search and browse television programme information. Internally, a context machine could be built relaying on public systems to deliver context (E.g. Wikipedia [6], IMDb [7], DBPedia [8], MusicBrainz [9] etc.) and provide a better view and navigation of programme information.
- *Product Marketplace / Wholesale Catalog* - A marketplace for service providers, partners, subsidiaries or divisions to search, browse and showcase content owners or broadcasters rich media products portfolio. Business users will be able to log into the system, search, browse, preview, buy or license media content products (E.g. television series, live sports offering, games, etc.). Various Open APIs could be provided to easy software integration with third parties.

9. CONCLUSION

In this paper we have provided an high level overview of the Derrick Metadata Platform, its vision and challenges for providing a next generation Internet platform for television metadata. The platform is under development at Derrick Media Ltd. and more information is available at http://www.derrickplatform.com.

10. REFERENCES

[1] Facebook - http://www.facebook.com

[2] Twitter - http://twitter.com

[3] "Descriptive Metadata for Television: An End-to-End Introduction", Mike Cox, Linda Tadic, Ellen Mulder - Focal Press

[4] "Being Digital" by Nicholas Negroponte - Vintage Books; 1st Vintage Ed edition (1 Jun 2000)

[5] "The Social Data Revolution(s)", by Andreas Weigend - http://blogs.hbr.org/now-new-next/2009/05/the-social-data-revolution.html

[6] Wikipedia - http://www.wikipedia.org/

[7] The Internet Movie Database (IMDb) - http://www.imdb.com

[8] DBPedia - http://dbpedia.org

[9] MusicBrainz - http://musicbrainz.org/

A Software Product Line Requirements and Reference Architecture for TV Navigation System

Antonio Ferreira dos Santos Jr.
Universidade Federal do Amazonas
Av. Gal. Rodrigo Octávio Jordão Ramos, nº 3000 – Bairro Aleixo – Manaus – AM – Brazil
+55 92 8826 1924
ajunior.analista@gmail.com

Vicente Ferreira de Lucena Jr.
Universidade Federal do Amazonas
Av. Gal. Rodrigo Octávio Jordão Ramos, nº 3000 – Bairro Aleixo – Manaus – AM – Brazil
+55 92 3305 6680
vicente@ufam.edu.br

ABSTRACT
In this paper we describe the goals, features, and reference architecture for a software product line of a TV navigation system, designed to create many products to be used in several parts of the world, using ISDB, DVB and ATSC standards and several languages. Using those goals, features and reference architecture, a manufacturer can develop several TV navigation systems at reasonable costs, enhanced quality and reduced time to market (advantages given by reusing a common base of technology), each with a similar interface, the same recommendation system and EPG, and one of the standards (ISDB, DVB or ATSC), one language for a low-end, mid-end or high-end set-top-box market segment.

Categories and Subject Descriptors
D.2.11 [**Software**]: Software Architectures – *Domain-specific architectures.*

General Terms
Documentation, Design

Keywords
TV Navigation System, Electronic Program Guide (EPG), Software Product Line, Reference Architecture.

1. INTRODUCTION
The industry, to reach more consumers, has adopted the mass customization concept to develop several and different products sharing the same components. In other words, launch similar products attending specific requirements from several market segments. The methodoloy used to develop software for mass customization is the software product line engineering (SPLE) [1,2], that uses component reuse concepts in an efficient and systematic way, resulting on several software products sharing a common plataform, which are part of a software product line (SPL) [1,2], developed in less time and cost and better quality compared to the development of those same several software products isolatedly. The SPLE has two main processes: domain engineering, which defines and realises the commonality and the variability of the software product line, and application engineering, which derives specific applications by exploiting the variability of the software product line.

The reference architecture is a core architecture that captures the highlevel design for the applications of the software product line. It is defined during the domain application engineering, and determines which software components are reusable and thus have to be developed.

In this paper, the reference architecture of a TV navigation system software product line is designed to derive products to be used in several parts of the world. Each product having ISDB, DVB or ATSC standard, one or more of several languages, and having specific characteristics for low-end, mid-end and high-end market segments. Using this reference architecture a manufacturer can, for example, deliver one TV navigation system configured for brazilians with the portuguese language and ISDB standard, other for americans with the english language and ATSC standard, among other configurations, both with a similar user interface and the same recommendation system, reusing several components.

The rest of this paper is organized as follows: Section 2 presents a brief introduction of the software product line techniques used to define the reference architecture; Section 3 presents the TV navigation system goals and features; Section 4 presents the reference architecture; and the Section 5 summarizes the paper.

2. SOFTWARE PRODUCT LINES
Mass customisation is the large-scale production of goods tailored to individual customers' needs. Platforms are any base of technologies on which other technologies or processes are built. The combination of mass customisation and a common platform allow us to reuse a common base of technology and to bring out products in close accordance with customers's wishes [1].

Software product lines are designed to provide customized products at reasonable costs, enhanced quality and reduced time to market. The costs are reduced when artefacts from the platform – which contains common artefacts and technological capabilities – are reused in several different kinds of system, this implies a cost reduction for each system. The quality enhancement comes from the platform artefacts reviewing and testing in many products. The time to market is initially higher, as the common artefacts have to be built first; after having passed this hurdle, the time to market is considerably shortened as the artefacts can be reused for each new product [1]. The cost and time to market is better after the third product, compared to the development of each product individually [2]. Several software product lines has been implemented, among those are Philips Consumer Electronics Software for Television, Nokia Mobile Phone, Nokia Networks and Bosch Gasoline System [3].

To develop an entire software product line two processes are used: the domain engineering and the application engineering. The former realises the commonality and the variability of the

product line. And the later builds the applications (individual products) of the product line by reusing artefacts from the domain engineering and exploiting the product line variability [1].

The domain requirements engineering sub-processes encompasses all activities for eliciting and documenting the common and variable requirements of the product line. Its output comprises reusable, textual and model-based requirements, defined on the section 3 as TV Navigation System Requirements, and, the variability model of the software product line, defined on the section 4 as the Reference Architecture. Both were provided following the domain requirements and domain design sub-processes, from the domain engineering process.

3. TV NAVIGATION SYSTEM REQUIREMENTS

A TV Navigation System combines program recommendation, sorting and retrieval to make it easier for the viewer to select programs based on various individual viewing habits [5], and is the guide of the viewer to select services and applications, initiate interoperable applications, boot loading, and store user profiles [9]. The goals, features and internal requirements are described in the following sub-sections.

Table 1. Low end Features

This Channel Information: The viewer can quickly determine important facts about this channel program/event like what's on now, next, tomorrow, what's on until 10 days afterwards [4]. Data is generated from a TV service provider for its service only. This feature provides [9]: * A channel info bar, which displays current channel and program/event information like: service/channel name, channel logo, event name, detailed information about the current event and the date and time of the set-top box, while the current program/event is playing; * A channel guide, which displays the present and the following program information (same as channel info bar) about current selected channel, while the current program/event is playing in a small screen; and * A program guide which displays a seven day program schedule information of the current channel, showing detailed information of the selected program/event.
Block Programs/Events: Supresses some not appropriated content for some viewers [4]. This feature blocks programs/events according to a list created and maintained by the viewer explicitly. The receiver can be programmed to not show specific channels or programs/events with violent or sex content, as the viewer can be children at that time.

3.1 Goals

Develop several TV Navigation System products based on a common platform. Those products will be embedded in several set-top-boxes or TVs of the same manufacturer with different characteristics, to be used in any part of the world for different market segments, and each can have from basic to sophisticated features.

3.2 Features

Tables 1, 2 and 3 present the end-user (viewer) visible characteristics of those TV Navigation Systems. Those characterictics (features) was defined through a research in EPG domain. The low-end set-top-box market segment has current (this) channel information and block program/event features (table 1), the mid-end has low-end, EPG schedule and reservation features (table 2), the high-end has mid-end, recommendation, PVR and second tuner features (table 3).

3.3 Internal Requirements

Table 4 presents mandatory and optional requirements for this TV Navigation System software product line not visible for the viewer, except for the User Interface. Those requirements were defined to obtain and store program/events information and viewer profile, as well as, define languages.

Table 2. Mid End Features

Search Programs/Events: Retrieves available programs according to a personal filter [5]. The function of search programs/events [4, 5, 8, 10] according to keywords provides a list of programs/events which match with a provided keyword. This keyword can be provided by the viewer through a real or virtual keyboard, remote control, or through a keyword list provided by the receiver to be chosen by the viewer. In this last case, this keywords list will be created from current programs/events available, and will be sorted according to a stored viewer profile [5].
Navigation on a Program Guide: The viewer selects or reserves programs/events for view and/or record them. Viewer navigates in an electronic program guide (EPG) showing rows representing channels and columns representing program/events with their names, beginning time and duration [4,5].
Reminders on Favorite Programs: prompts the viewer when favorite or defined programs/events are on [8]. The favorites will be prompt based on viewer's personal profile, if the user turn this option on; and the defined will be prompted if the viewer reserved it before, through the EPG.

4. TV NAVIGATION SYSTEM REFERENCE ARCHITECTURE

The reference architecture is a core architecture that captures the high level design for the applications of the software product line. It includes the variation points and variants documented in the variability model realized by components [1]. The figure 1 presents the component diagram representing all navigation system components (which implements the features) and the variability model, in which there are 3 variation points (represented by triangles): language, market segment and standard. Each of those is explained bellow, where the tables (1,2 or 3) –describing the features implemented by each variant (represented by rectangles)– are defined in parenthesis. The *language* variation point has variants representing languages for the TV Navigation System, depending on the countries where it will be used. Although some languages were specified in the variability model, other languages can be added. The set-top-box requires at least one of those languages.

The *market segment (MS)* variation point has variants representing low-end, mid-end and high-end market segments. The arrows pointing to the component diagram indicates the required components that implements those variants.

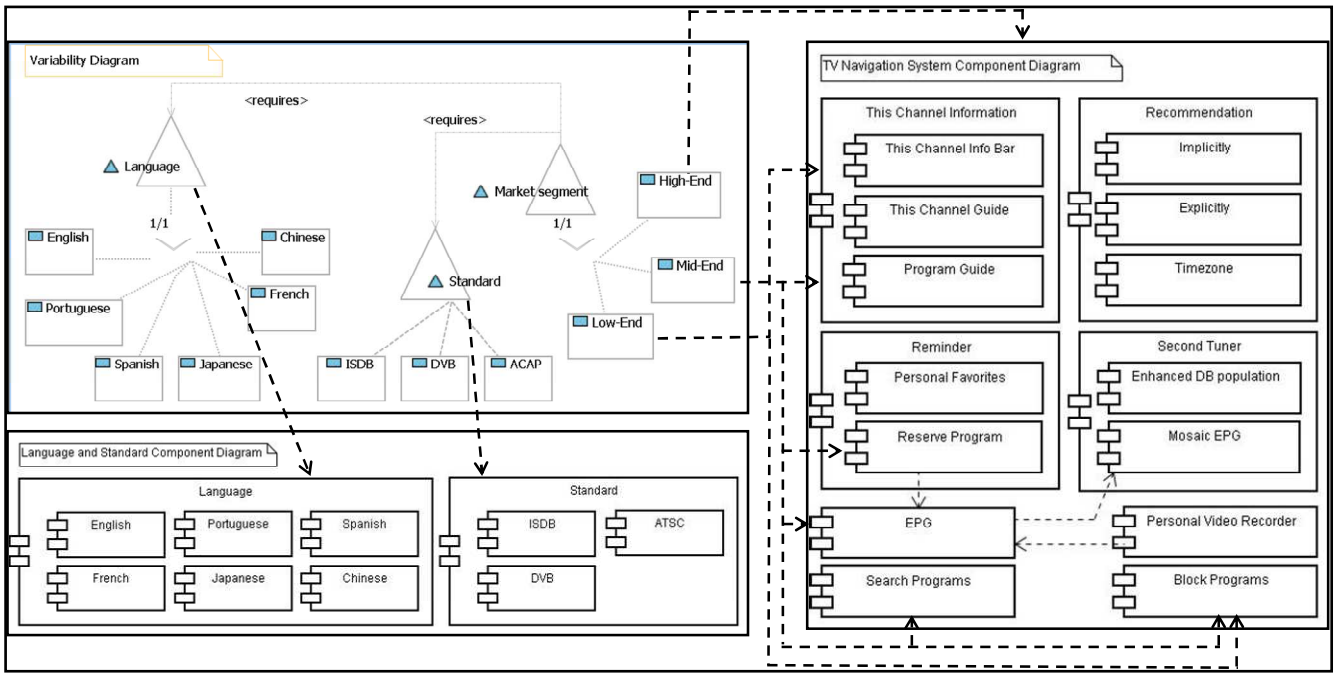


Figure 1. Variability and Component Diagrams (reference architecture)

The low-end variant (table 1) implements the *this channel information* component. The mid-end variant (tables 1 and 2) is implemented by the following components: *this channel information, reminder, block program/events and search program/events*.

Table 3. High End Features

Personalized Recommendation System: The personalized recommendation system [5,10,11,12] recommends content according to viewer's preferences and interests, given implicitly, where the system will track user preferences , or explicitly, where the viewer can choose the general interests of genre, channel or rating as the explicit profile. As output, the system can simply select a channel automaticaly, provide a list of programs/events to be selected by the viewer, or provide an EPG schedule sorted by most suitable program/events. The recommendation can vary depending on time zone (morning, afternoon, evening or night). In the morning, for example, a viewer might want to check out the news and the weather report for that day; while in the evening, the same viewer might like to watch a more dramatic or entertaining program such as a movie, drama, or live sports event [5].
Personal Video Recorder (PVR): Provides functions such as fast forward, rewind, pause, time shift the broadcast, skip over commercial, automatic program recording, instant replay in live programs, guiding the plot and choosing camera angle [6, 8].
Second Tuner: Although it has picture in picture (PiP), and other applications, it can also be applied on: Personalized mosaic EPG which shows a screen with 9 channels, updating their video periodicaly using the first tuner while presenting current audio and video of the selected channel [7] ; and on improving the way to populate Service Information Database using the second tuner to receive EPG information while the viewer watch a TV event demodulated from the first tuner.

The high-end variant (tables 1, 2 and 3) is implemented by all components in TV Navigation System Component Diagram (Figure 1). Each set-top-box requires only one market segment.

Table 4. Internal Requirements

Viewer Profile Engine: Collect and store information about the viewer in an implicity or explicity way. In the implicity way the viewer behavior is collected and stored when he/she is watching TV. In the explicity way, the viewer informs the channels and the kind of program/event genres that he/she likes according to his/her preference through a user interface.
Service Information Engine: Receive service information (SI) from Transport Stream and other sources, and stores it into the service information database.
Profile Database: Stores information about viewer's preferences given implicitly or explicitly. This information will be used for program/event recommendation.
Service Information Database: Stores TV broadcast schedule for searching, filtering, sorting and selecting programs/events [5, 10].
Multiple language support: As this TV Navigation System is designed to be used in any part of the world, its user interface must support the target countries language.
User Interface: Must be easy to use and navigate. All products will have similar user interface
Standards: Each product will receive and parse the broadcast signal, through the Service Information Engine, using the ISDB [13] , DVB [14] or ATSC [15] standard.

The *standard* variation point has variants representing the ISDB [13], DVB [14] and ATSC [15] DTV standards, which are implemented by the components with the same name, described in

Language and Standard Component Diagram (Figure 1). Each set-top-box requires only one standard.

Four examples of individual products based on this software product line of TV Navigation system reference architecture are the TV navigation systems A, B, C and D, specified bellow:

Product A: this product is designed for the american high-end market segment. The language is english, the standard is ATSC, the features are all defined in tables 1, 2 and 3.

Product B: this product is designed for the european mid-end market segment. The language is english, the standard is DVB, the features are: this channel information, reminder on products/events, EPG, block and search programs/events.

Product C: this product is designed for the brazilian low-end market segment. The language is portuguese, the standard is ISDB, the features are this channel information and block programs.

Product D: this product is designed for the mexican low-end market segment. The language is spanish, the standard is ATSC, the features are this channel information and block programs.

5. CONCLUSION

In this paper, we described the reference architecture of a TV navigation system software product line (SPL) designed to implement several TV navigation systems for three set-top-box market segments (low-end, mid-end and high-end), three DTV standards (ISDB, DVB and ATSC) and several languages. The low-end set-top-box market segment has this channel information and block program/event features; the mid-end has low-end, EPG schedule and reservation features; the high-end has mid-end, recommendation, PVR and second tuner.

This SPL will derive several TV navigation systems that share common requirements, improving quality and reducing time to market and costs on the whole products development. Those features were defined researching several papers on TV navigation system; and the reference architecture was designed following the SPLE domain engineering process.

This reference architecture is already being used to implement three TV navigation systems for brazilians: one for low-end, other for mid-end, and a third for high-end market segment, with portuguese language and ISDB standard. There are several papers on instances of SPLs subject, and also several on TV navigation systems, but no papers on SPLs applied to TV navigation system subject where found to be compared to our proposal.

6. ACKNOWLEDGMENTS

Our thanks to FAPEAM – Amazonas State Research Support Foundation – for providing the scholarship during the research work; to CNPq – The National Council for Scientific and Technological Development – for providing resources to our lab; and to Ricardo Rosa and Leonard Costa for reviewing this paper.

7. REFERENCES

[1] Pohl, K., Böcke, G., Linder, F. 2005. Software Product Line Engineering – Foundations, Principles and Techniques. Springer-Verlag Berlin hedelberg. Germany.

[2] Northrop, L., Clements, P. 2007. A Framework for Software Product Line Practice, Version 5.0. http://www.sei.cmu.edu/productlines/framework.html.

[3] Linden, F., Schmid, K., Rommes, E. 2007. Software Product Lines in Action: the Best Industrial Practice. In Product Line Engineering. Springer-Verlag New York, Inc.

[4] Basic, R., Mocinic, M. 2002. User's requirements for electronic program guide (EPG) in interactive television (iTV). In Video/Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom (16-19 June 2002). 457-462. DOI=10.1109/VIPROM.2002.1026702.

[5] Isobe, T., Fujiwara, M., Kaneta, H., Uratani, N., Morita, T. 2003. Development and features of a TV navigation system. In IEEE Transactions on Consumer Electronics.Vol. 49, Issue 4, Nov, 2003. 1035–1042. DOI = 10.1109/TCE.2003.1261192.

[6] Zhang, Z., Li, J., Pan, L., Ni, H. 2006. Implementation of an advanced TV Navigation System. In Proc. of the 8th Int. Conference on Signal Processing. ICSP 2006. Vol. 4. 16-20 Nov. 2006. DOI = 10.1109/ICOSP.2006.346097.

[7] Han, J. Han, I., Park, K. 2008. User-Configurable Personalized Mosaic Electronic Program Guide. In IEEE Transactions on Consumer Electronics. Vol. 54, Issue 1, Feb, 2008.192–196. DOI= 10.1109/TCE.2008.4470043

[8] Jensen, J. 2005. Interactive television: new genres, new format, new content. In Proc. of the second Australasian conference on Interactive entertainment (Nov, 2005). 89-96.

[9] Peng, C., Lugmayr, A., Vuorimaa, P. 2002. A Digital Television Navigator. In Multimedia Tools and Applications. Volume 17, Number 1. May, 2002. 121-141. DOI= 10.1023/A:1014687823960

[10] Zhang, H., Zheng, Z., Yuan, J. 2005. A personalized TV guide system compliant with MHP. In IEEE Transactions on Consumer Electronics. Volume 51, Issue 2, May 2005. 731-737. DOI= 10.1109/TCE.2005.1468026

[11] Ehrmantraut, M., Härder, T., Wittig, H., Steinmetz, R. 1996. The personal electronic program guide—towards the pre-selection of individual TV programs. In Proceedings of the fifth international conference on Information and knowledge management – CIKM ’96. 243-250.

[12] Weiß, D., Scheuerer, J., Wenleder, M., Erk, A., Gülbahar, M., Linnhoff-Popien, C. 2008. A user profile-based personalization system for digital multimedia content. In Proc. of the 3rd int. conference on Digital Interactive Media in Entertainment and Arts – DIMEA ’08. 281-288.

[13] ABNT NBR15603-2. 2007. “Televisão digital terrestre – Multiplexação e serviços de informação (SI). Parte 2: Estrutura de dados e definições da informação básica de SI”. The Brazilian Specification for System Information.

[14] European Telecommunication Standard Institute (ETSI) 300 468. 1997. “Digital Video Broadcasting (DVB); Specification for Service Information (SI)”.

[15] Advanced Television Systems Committee (ATSC) A/65. 1997. “Program and System Information Protocol for Terrestrial Broadcast and Cable (PSIP)”.

Evaluation of a TV Programs Recommendation using the EPG and Viewer's Log Data

Kazuki Ikawa
Graduate School of Engineering,
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo, Japan

Tomohiro Fukuhara
National Institute of Advanced Industrial Science and Technology
2-3-26 Aomi, Koto-ku
Tokyo, Japan
tomohiro.fukuhara@aist.go.jp

Hideki Fujii
Graduate School of Engineering,
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo, Japan
fujii@save.sys.t.u-tokyo.ac.jp

Hideaki Takeda
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku
Tokyo, Japan
takeda@nii.ac.jp

ABSTRACT

Evaluation results of a TV program recommendation using EPG (Electronic Program Guide) and viewer's log data are described. Two experiments using the log data of a Japanese video service provider are conducted: (1) an experiment of prediction of TV programs that each viewer usually watched, and (2) a survey of TV program recommendation using the log data and a questionnaire. High precision values are obtained in the first experiment using *keyword-based* and *celebrity-based* recommendation methods for terrestrial and satellite broadcasting programs. Meanwhile, the *channel & time-based* recommendation method for programs in paid channels of the video service provider obtained the high precision in the second experiment. An overview of experiments and results are described.

Categories and Subject Descriptors

H.4.2 [**Information Systems Applications**]: Types of Systems

General Terms

EXPERIMENTATION

Keywords

TV program recommendation, evaluation

.

1. INTRODUCTION

Today, many TV programs are broadcasted on various channels. Terrestrial digital broadcasting services and satellite broadcasting services brought us many channels and programs. As the increase of channels and programs, recommendation services for TV programs and channels are needed that allow us to choose suitable programs to watch for each use[4]. For designing a good TV program recommendation service, fundamental data that compares performances among various recommendation methods is needed[6].

We had two experiments of a TV program recommendation using EPG (electronic program guide) and the viewer's log data: (1) an experiment of prediction of TV programs that each viewer watches usually, and (2) a survey of TV program recommendation using the log data and a questionnaire. We created a prototype system of a TV program recommender system that recommends programs for each user using EPG and the viewer's log data. An overview of experiments and results are described.

In the following, section 2 describes the related work. Section 3 describes an overview of our approach for TV program recommendation. Section 4 describes an overview of our experiment and its results. Section 5 discusses the results of the experiments. We summarize arguments in section 6.

2. RELATED WORK

There are many work on TV program recommendation. TiVo, which is a TV program recommendation system based on the collaborative filtering (CF) method, is a popular product in U.S[1]. Although the CF method works well where a large number of judgement data is available, we consider that the CF method have difficulties when new programs that few people watched before.

Smyth and Cotter introduces an overview and performance of a personalized TV guide system[8]. They also adopted the CF method. In their experiment, 310 households attended to evaluate the precision of recommendation, and found that

Table 1: An example of a viewer's log (dummy data)

Household ID	Date	Time	Channel ID
1000	01/Nov/2009	07:18	235
1000	01/Nov/2009	07:19	235
1000	01/Nov/2009	07:20	234

Table 2: An example of EPG data.

Channel ID	Start Time	End Time	Title	Description
235	01/Nov/2009 07:00:00	01/Nov/2009 07:59:00	***	***

61% of households answered positively. Their result showed that the CF method showed the best performance, compared to the content-based (CB) recommendation and a naive random recommendation.

Although there exist recommender systems based on the CF method, few data is reported with respect to the evaluaion of the CB recommendation. In this paper, we will evaluate the performance of a CB recommendation method using the real data of a video service provider.

3. APPROACH FOR THE TV PROGRAM RECOMMENDATION

3.1 Log data

We used the viewer's log data of a video service provider in Japan. We call this provider *TV-999* in this paper. Each household has a set top box (STB) that receives TV programs of *TV-999*. *TV-999* broadcasts terrestrial programs, satellite broadcasting (BS[1]) programs, and *paid channels* such as movie, drama, sports, animation and so on[2]. The view history of a household is recorded at the server machine of the provider[3].

Table 1 shows an example of log data. The log data consits of following parts: (1) *household ID*, (2) *date and time* when the household watched a channel, and (3) *Channel ID* at which the household watched. The data is recorded in each minute. In this case, household ID 1000 watched channel ID 235 and 234 during 7:18 through 7:20 a.m. on November 1st, 2009. For zapping action, we removed the view history of which duration is under five minutes.

3.2 EPG data

EPG data contains *content ID*, *date and time* when the content is broadcasted, *duration* of the content (minute), *title* and *summary* of the content. Table 2 shows an example of EPG data. We extract keywords and names of celebrities appeared in title and description. We can identify which household watched which programs by combining EPG and the log data[5].

[1]BS is a common name of one of the direct broadcast satellites in Japan.
[2]There are 30+ paid channels in *TV-999*.
[3]The log data is recorded under the agreement of each household. If a household does not want, the data is not recorded.

3.3 Recommendation method

Based on programs that each household watched, we first calculate feature values of keywords and celebrities using following formulas[2][3][7].

$$P_{key}(k) = \frac{|Watched(k)|}{|Programs(k)|} \quad (1)$$

$$P_{person}(p) = \frac{|Watched(p)|}{|Programs(p)|} \quad (2)$$

$P_{key}(k)$ is a function that calculates a feature value for a keyword k. $Watched(k)$ is a set of programs that a household watched. $Programs(k)$ is a set of programs that contains the keyword k in their EPG data. $P_{person}(p)$ is a function that calcuates feature value for a celebrity person p. $P_{key}(k)$ and $P_{person}(p)$ are calculated for each household.

For calculating recommendation values of a program *title*, we use following formulas[9].

$$R_{key}(title) = \sum_{k \in Keywords(title)} P_{key}(k) \quad (3)$$

$$R_{person}(title) = \sum_{k \in Persons(title)} P_{person}(p) \quad (4)$$

$$R_{mix}(title) = \frac{R_{key}(title)}{\max_{t \in TITLES} R_{key}(t)} + \frac{R_{person}(title)}{\max_{t \in TITLES} R_{person}(t)} \quad (5)$$

$R_{key}(title)$ is a function that calculates a recommendation value for a program *title* containing a keyword k in its EPG data. $Keywords(title)$ is a function that returns a set of keywords contained in the EPG data of a program *title*. Formula (3) is used as a *keyword-based* recommendation in the next section.

$R_{person}(title)$ is a function that calculates a recommendation value for a celebrity *person* appeared in the EPG of program *title*. $PERSONS$ is a set of all of persons appeared in all of EPG data. Formula (4) is used as a *celebrity-based* recommendation.

$R_{mix}(title)$ is a function that calculates recommendation value for a program *title*. $TITLES$ is a set of all titles appeared in all of EPG data. Formula (5) is used as a *keyword & celebrity-based* recommendation.

4. EXPERIMENTS

4.1 Prediction of daily viewing programs

Approach

We evaluated the accuracy of prediction of programs that each household daily watches. We used the log data in February 2008. The data is divided into twofold: (1) the first part (February 1st to 15th 2008) is the *training data* which is used for calculating feature values of keywords and celebrities. (2) the second part (February 16th to 29th 2008) is the *test data* that is used for evaluating prediction of programs.

The log data contains the view history of 923 households. 111,343 programs are contained in this data. Among these programs, 29,887(26.8%) are terrestrial/BS programs, and 81,456(73.2%) are programs of paid channels.

With respect to programs watched by households, 669 programs are watched by each household in the test stage. Among these programs, 70.1% (469) are terrestrial / BS programs, and 29.9% (200) are programs of paid channels.

We prepared two types of recommendations: (1) *keyword-based* recommendation, and (2) *celebrity-based* recommendation. For the first type of recommendation, we use the formula (3) to calculate a recommendation value of a program for a household. For the second type of recommendation, we use the the formula (4) to calculate recommendation a value of a program for a household. We randomly picked up 10 households, and evaluated 2*100 programs, which is created by two types of recommendations, for each household. We used terrestrial / BS programs because the large amount of log data was available for these programs.

As evaluation measures, we used precision and recall measures described in the following.

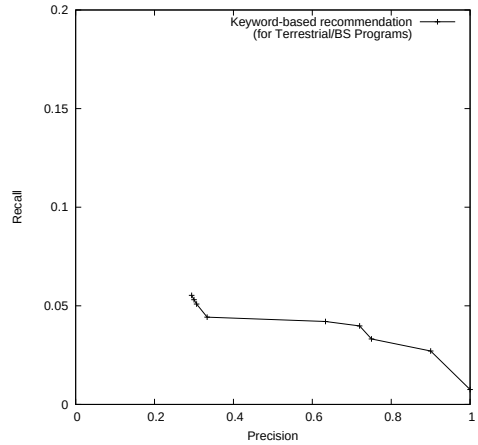
$$Precision = R/N \quad (6)$$

$$Recall = R/C \quad (7)$$

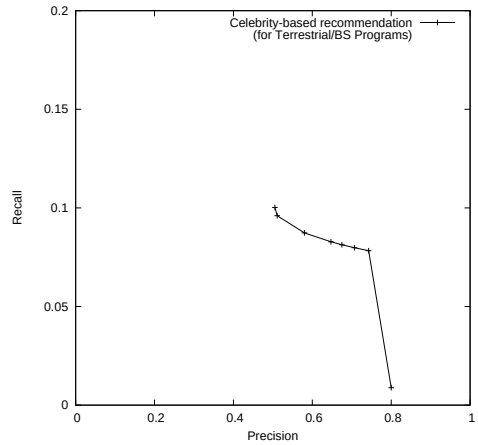
R is the number of programs that the system suggested and a household watched, and N is the number of programs that the system suggested. C is the number of programs that a household watched.

Results

Figure 1(a) and Figure 1(b) show the results for *keyword-based* and *celebrity-based* recommendations respectively. In these figures, the recall values are quite low because the number of C in the formula (7) is much larger than the number of programs that we recommended in the experiment[4]. So, the recall values are low. With respect to the precision, both of recommendation types showed high precision values. We considered that predicting programs that a household usually watches based on *keywords-based* and *celebrities-based* recommendation was possible.



(a) Recall and precision of the keyword-based recommendation for terrestrial/BS programs.



(b) Recall and precision of the celebrity-based recommendation for terrestrial/BS programs.

4.2 Evaluation of recommendations using a questionnaire

Approach

Based on the results of the first experiment, we had a survey of TV program recommendation using keywords and celebrities. We prepared a questionnaire containing recommended programs for households using the log data. In this experiment, we chose programs in paid channels.

Table 3 shows the periods of this experiment. As a training period, we used the log data during November 9th through 22nd (14 days), 2009. As a test period in which TV programs are recommended in the questionnaire, we chose the period during November 28th to December 4th (7 days). For avoiding the change of preferences of each household[5], we sent the questionnaire after the test period, i.e., each household received the questionnaire after December 5th.

The questionnaire contains programs recommeded by *keyword-based* and *celebrity-based* recommendations for each household. We asked each household up to five programs for each day, i.e., up to 35 programs broadcasted in the test period are asked. We asked each household to judge programs according following four categories: (1) *Watched & Good*, (2) *Watched & ¬Good*, (3) *¬Watched & Expected*, and (4) *¬Watched & ¬Expected*. Each household checked one of these categories for each recommendation.

For comapring performances of recommendation types, we classified households into four types. Table 4 shows the list of recommendation types. Type $A\&B$ is the mixture of type A and B, which is calculated using the formula (5). Type C is the baseline type where programs are chosen randomly according to the joint probability of channels and time-slots that a household usually watches TV. This probability is calculated from the log data of a household in the training period. We used the log data of 907 households of the service. We sent a questionnaire containing recommendation of programs for each household. Questionnaires are sent to 875 households via mail on December 4th, 2009[6]. By December 25th, 2009, 605 answers (69.3%) are returned.

[4]We only evaluated 100 programs for each household

[5]This was a request from *TV-999*.

[6]Households received the questionnaire after December 5th, 2009.

Table 3: Periods of experiment.

Dataset	Term	# of Days
Training	Nov.9th - Nov.22nd, 2009	14 days
Test	Nov.28th - Dec.4th, 2009	7 days

Table 4: Types of recommendation.

Rec. Type	Description	# of Received/ Recommended (Ratio)
A	*Keyword-based*	112/160 (70.0%)
B	*Celebrity-based*	141/200 (70.5%)
A&B	*Keyword&Celebrity*	132/200 (66.0%)
C	*Channel&Time*	220/315 (69.8%)

Results

Table 5 shows the summary of the result. Most of recommendations were judged as ¬*Watched & ¬Expected*. Except for ¬*Watched & ¬Expected* category, type *C* obtained the best score in *Watched & Good* category. This was an unexpected result.

Table 6 shows another summary of the result which is organized according to viewpoints of *prediction* and *recommendation*. The *prediction* viewpoint is the union of *Watched & Good* and *Watched & ¬Good* evaluations, and *recommendation* viewpoint is the union of *Watched & Good* and ¬*Watched & Expected* evaluations. For both of viewpoints, type *C* (*channel & time*) showed the best score. Whereas type *B* showed the lowest score in both of *prediction* and *recommendation* viewpoints. Althgouh type *A* at the *recommendation* viewpoint showed the second score, it did not exceed the type *C*.

5. DISCUSSION

One of our unexpected things is that the *keyword-based* and *celebrity-based* recommendations did not work in the second experiment. Although we expected that both of these recommendations would show good scores, the results was a contrary. Especially, type *celebrity-based* recommendation was the worst among four recommendations in Table 6.

For reasons of this failure, we consider that there are preferences for paid channels. Because paid channels in *TV-999* are highly specialized, we consider that each household already has its own preference for specific channels. For understanding this phenomenon, further investigation using the qualitative data such as the age of a respondent, number of people in his/her household, and channels and celebrities that s/he likes is needed.

6. CONCLUSION

We described evaluation results of experiments on TV program recommendation using EPG and viewer's log data. We had two experiments: (1) a TV program prediction experiment using the log data, and (2) a survey of TV program recommendations using the log data and a questionnaire. In the latter experiment, we found that the proposed methods did not work well, and the baseline method recommendation showed the best score. We will continue to investigate this phenomenon by analyzing the obtained questionnaire data.

Table 5: Results of the experiment (%).

Rec. Type	Watched& Good	Watched& ¬Good	¬Watched& Expected	¬Watched& ¬Expected
A	9.9	0.5	17.5	72.1
B	3.1	0.4	12.1	84.4
A&B	6.3	0.4	16.7	76.6
C	13.9	1.1	16.3	68.7

Table 6: Performances with respect to prediction and recommendation.

Rec. Type	*Prediction* (*Watched&Good* ∪ *Watched&¬Good*)	*Recommendation* (*Watched&Good* ∪ ¬*Watched&Expected*)
A	10.4 (%)	27.4 (%)
B	3.5 (%)	15.2 (%)
A&B	6.7 (%)	22.9 (%)
C	**15.0 (%)**	**30.2 (%)**

7. REFERENCES

[1] K. Ali and W. van Stam. Tivo: making show recommendations using a distributed collaborative filtering architecture. In *KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 394–401. ACM, 2004.

[2] L. Ardissono, C. Gena, P. Torasso, F. Bellifemine, A. Chiarotto, A. Difino, and B. Negro. Personalized recommendation of TV programs. In *AI*IA 2003: Advances in Artificial Intelligence, Lecture Notes in Computer Science*, volume 2829, pages 474–486, 2003.

[3] A. Bär, A. Berger, S. Egger, and R. Schatz. A lightweight mobile TV recommender. In *EUROITV '08: Proceedings of the 6th European conference on Changing Television Environments*, pages 143–147. Springer-Verlag, 2008.

[4] A. Basso, M. Milanesio, and G. Ruffo. Events discovery for personal video recorders. In *EuroITV '09: Proceedings of the seventh european conference on European interactive television conference*, pages 171–174. ACM, 2009.

[5] G. Graefe. Query evaluation techniques for large databases. *ACM Comput. Surv.*, 25(2):73–169, 1993.

[6] M. Gude, S. M. Grünvogel, and A. Pütz. Predicting future user behaviour in interactive live TV. In *EUROITV '08: Proceedings of the 6th European conference on Changing Television Environments*, pages 117–121. Springer-Verlag, 2008.

[7] R. J. Mooney and L. Roy. Content-based book recommending using learning for text categorization. In *DL '00: Proceedings of the fifth ACM conference on Digital libraries*, pages 195–204. ACM, 2000.

[8] B. Smyth and P. Cotter. A personalized television listings service. *Commun. ACM*, 43(8):107–111, 2000.

[9] W. E. Spangler, M. Gal-Or, and J. H. May. Using data mining to profile TV viewers. *Commun. ACM*, 46(12):66–72, 2003.

YouToRemix : the online audiovisual quotting live remix application project

Milena Szafir
Manifesto21.TV
www.manifesto21.com.br
São Paulo-SP, BRAZIL

milena@manifesto21.com.br

ABSTRACT

Nowadays it has become clear that processes of digitalization and convergence have transferred audiovisual production from the television screen to the Internet environment. How can one use a video as a reference to be remixed without having to resort to complicated (and heavy) non-linear video applications? As a consequence of this current state of affairs, open issues concerning this topic will be here discussed and analyzed.

This article wishes to show a different use of audiovisual content through online video platforms. If we consider Youtube (by Google) as an example of a kind of video-library – online audiovisual archives – let us link these databases to one another and mix them, creating a new video, an audiovisual rhetoric. Then, the user will be able to take excerpts from one video and link (mix) it to other ones in real time. Aiming at organizing the online audiovisual content like quotes, this project (an online video-library live remix) is part of my master degree studies and is currently being developed. Does this kind of project interest someone else? The aim of this paper is to give a simple overview of the possible new ways to make online videos with the focus on users' needs and experiences.

Categories and Subject Descriptors

D.2.2 [**Software**]: Design Tools and Techniques – *Computer-aided software engineering (CASE), User Interfaces* | H.5.4 [**Information System**]: Hypertext/Hypermedia (I.7, J.7) – *Theory, User Issues* | J.5 [**Computer Applications**]: Arts and Humanities – *Arts, fine and performing*

General Terms

Documentation, Performance, Design, Experimentation, Theory.

Keywords

remix, mashup, YouTube, Google, real time, live, vj'ing, participatory, online, web, sharing, Youtubteca, Youtubemix, Remixpress

1. INTRODUCTION

(motivation for app development and its main usability)

In this paper I would like to discuss how we should design and evaluate interactive audiovisual content applications, trying to highlight some relevant aspects of web-sharing-video issues, like "the remix", toward a new participatory use of online video platforms, in particular YouTube.

Online video represents a new challenge on how to link, create and present content that maximizes the internet user experience. During this first decade of XXI century, research often focused on several aspects of this new audiovisual online participatory scheme, but neglected aspects that might be of interest when trying to understand the wishes of a new generation inside a digital culture and its available technologies. A new concept for the content production has been created to be used by ex-spectators.

Some years ago, blogs (i.e. "web log") made a revolution on the Internet, and suddenly anyone was able to express himself with total freedom – first by words and later by pictures followed by embedded videos –, the wordpress came then in a open source structure where any user could assemble his own aesthetic website as well as its usability:

> *"WordPress is an open source CMS, often used as a blog publishing application powered by PHP and MySQL. It has many features including a plugin architecture and a templating system. ... WordPress has a templating system, which includes widgets that can be rearranged without editing PHP or HTML code, as well as themes that can be installed and switched between. The PHP and HTML code in themes can also be edited for more advanced customizations. WordPress also features integrated link management; a search engine-friendly, clean permalink structure; the ability to assign nested, multiple categories to articles; and support for tagging of posts and articles."*[1]

> *"WordPress is a state-of-the-art publishing platform with a focus on aesthetics, web standards, and usability. WordPress is both free and priceless at the same time. More simply, WordPress is what you use when you want to work with your blogging software, not fight it."*[2]

Hence, throughout personal diaries, blogs turned into a good way to organize online database for their authors (and his readers). A database organization by tags and its reference (or quotes) links. In a blog you do not loose your work in progress: it works like a diary, or rather a card index, where you can turn back to the research and its notes folowing a determined path by key-words. Nothing is lost in a place where every character can be simply and quickly remixed, attributing a new value to the network, making it into a "new thing".

The academic research practice, in a brief conceptualization, can be thought of as an audiovisual remix practice: we make our research from authors, journals, images, offline (and online) database, etc; then we analyze these data, remix and transform them into a 'new' value; a thesis. But the main difference between these two practices is a methodological one, which, the way I see it, gives this short article its relevance to what I originally called Youtuboteca and later dubbed RemixPress. The existence of footnotes in audiovisual remix practice is quite unusual; we could mention, for example, some films in the 70's and television programs in the 80's (such as Debord and Godard's *Histoire(s) du Cinéma*, respectively, or even Masagão in Brasil with a movie called *Nós que aqui estamos por vós esperamos* 90's and several digital audiovisual remix works at beginning of our XXI century like *Rebirth of a Nation*, *The Society of Spectacle – digital remix* (made by american djs Spooky and Rabbi, respectively) and other ones by groups like *GNN* (EUA), *Media Sana* or by girls from *mm não é confete* (Brasil).

So, my application has been developed bearing in mind a methodological organization applied to the creation of an audiovisual remix practice for academic purposes.

2. ARTICLE'S STRUCTURE

(a)A brief contextualization of shared online audiovisual (the YouTube case); (b)A succinct statement about the current methodological work; (c)A defense of an audiovisual writing; (d)A description of the new application's references and some of its functions, as well as possible paths that could be followed.

3. YOUTUBE AS AN ONLINE AUDIOVISUAL PUBLIC LIBRARY

("2009 was a blistering year of growth in online video consumption."[3]

According to comScore[4], each user watches on average 81 videos per month on the Internet, and a little more than 100 million people accessed YouTube in January 2009.

> *"an analysis of worldwide video streaming activity from YouTube.com, confirming that an average of 100 million video streams were served per day in July 2006. ... On a daily basis, the site attracted an average of 6.2 million visitors worldwide, with 1.6 million residing in the U.S. ... In July 2006, YouTube served nearly 3 billion video streams worldwide, with slightly less than one-quarter of the total activity streamed to U.S. locations. On an average daily basis for the month, 96 million streams were served worldwide, and 21 million in the U.S."* [4]

YouTube – created in February 2005 by two ex-employees of eBay – reaches a surprising 100 million video views in July 2006 (representing 42.2% of the videos views on Internet that time)[5]. In the same year, around 65 thousand new digital videos were posted daily by common users. After some months, "Time Magazine" dubbed YouTube 'Invention of the Year'[6]. In October 2006, Google acquired YouTube for 1.65 billion dollars[7].

> *"What is revolutionary about YouTube is that it represents, in Pierre Levy's words, a 'natural and sound appropriation of discourse'; a website where Mass Media is quoted and remixed; where homemade media achieves public access and several subcultures produce and share media…The rhetoric of the digital revolution predicted that the new media would replace the old one; however, YouTube exemplifies a cultural convergence…the business model of YouTube creates added value by means of circulation…Although much or the Remix Culture is based on parody, this genre intensifies the emotional experience of the original material, bringing us deeper into the main characters' thoughts and emotions…The current status of YouTube makes it an inevitable platform for broadcasting content generated by its users…YouTube or not to YouTube, that is the question."*[8]

Hence, this videographic platform of a 'real time' dynamic flux can be perceived as a public archive of audiovisual memory from a Jean-Paul Fargier – related to Godard – writing:

> *"Television, since in its origins has been a device for the endless reproduction of the present and has a memory capable of boundless storage…one you can refer to not just as a testimony of the past but also to replace an impossible live image…it is therefore, a databank of all images, including those of the cinema."* [9]

So, YouTube becomes a global reference for on-demand video, and it is based on this concept that I consider this online platform as a main example for quoting, creation and aesthetic possibility of a networked audiovisual. It is beyond the scope of this work then, to talk about other different aspects of YouTube related to a "digital culture" such as its political and economical implications, etc.

> *"Videos viewed at YouTube surged from 6.3 billion in Jan '09 to 13.2 billion in Dec '09. To put this in perspective, Google closed its acquisition in Nov '06. In Jan '07 (the first month comScore publicly released online video data), YouTube notched 1.2 billion views."*[10]

> *"When the width of the band was finally expanded, allowing the broadcasting of videos along with text and music, the "adequate" YouTube was born, with the slogan Broadcast Yourself, its simple usability and its ambition for vertical growth in the shortest space of time, from "1" to "Global"… [thus] How could YouTube be exploited with scientific, journalistic and creative purposes of self-expression, communication and file-sharing?...YouTube allows anyone to perform their own "Bardic functions"*[11]

At first, the platform – based on Flash (property application by Macromedia, nowadays Adobe) – allowed just a simple and free way to publish and store videos up to a duration of 10 minutes. You could use text to comment on the videos and embed those in several blogs and websites, the only rule was (and still is): "Make Sure Your Video Does Not Infringe Someone Else's Copyrights"[12]. This, however, has not prevented people from sharing assorted audiovisual productions – old movies excerpts,

television ads and programs, dvds, etc – and, as the pioneer *Napster*, users feel free to share a huge amount of digital video data.

Consequently, my research is based on this daily breaking of the rule by millions of users, i.e. because of this infringement I consider YouTube, nowadays, above all an online audiovisual library. This platform is the online place where I look for videographic material, for the theme of my master's degree research, to construct what I have called *Audiovisual Rhetorics*.

> *"YouTube seems to provide an inexhaustible content generated by its users. But this very abundance (McCracken, 1998) might discourage us to question materials which are not found there. Historically the DIY movement aimed at enabling groups that had no commercial penetration to tell their own stories. If we admit that YouTube operates with no previous history, then we would be giving up what we have been fighting for and end up with much less than expected"*[8]

4. REMIX: A SUCCINT APRESENTATION

("unlimited remix" as the artistic and political technique of our time.[1])

According to Manovich[13], the remix practice became popular at the beginning of XXI century, creating a whole concept of a "new culture" in progress:

> *"Remixing originally had a precise and a narrow meaning that gradually became diffused. Although precedents of remixing can be found earlier, it was the introduction of multi-track mixers that made remixing a standard practice. With each element of a song – vocals, drums, etc. – available for separate manipulation, it became possible to "re-mix" the song: change the volume of some tracks or substitute new tracks for the old ounces.* Gradually the term became more and more broad, today referring to any reworking of already existing cultural work(s)*." (my highlights)*

The paradox that Manovich indicates in his text (related to copyright issues) is also revealed by Lessig (be showed forward). Why is the remix considered a violation in the audiovisual area?

> *"Yet we are left with an interesting paradox: while in the realm of commercial music remixing is officially accepted[2], in other cultural areas it is seen as violating the copyright and therefore as stealing."*[13]

But Manovich and Lessig disagree about quoting. Is quoting equivalent to remix practice? Let us bear with the former who wrote about remix and quoting; in his view they have different natures:

> *"The other older term commonly used across media is "quoting" but I see it as describing a very different logic than remixing.* ***If remixing implies systematically rearranging the whole text, quoting refers inserting some fragments from old text(s) into the new one.*** *Thus I think we should not see quoting as a historical precedent for remixing. Rather, we can think of it as a precedent for another new practice of authorship practice that, like remixing, was made possible by electronic and digital technology – sampling."*[13]

So, according to Eduardo Navas, "the remix is in the end ... a rearrangement of something already recognizable; ... In brief, the remix when extended as a cultural practice is a second mix of something pre-existent; the material that is mixed at least for a second time must be recognized otherwise it could be misunderstood as something new, and it would become plagiarism. Without a history, the remix cannot be Remix."[14]

4.1. The digital handicraft: a set of working methods

Nowadays, my work is done in a kind of handicraft way, i.e. I need to download (using free software easily found on the Internet) each and every video found, then collect them offline so that they can be worked on and maybe return this remix to the web. Moreover, this kind of digital handicraft with online videos is hard work because YouTube material is available in a .flv (Flash's codec) format, which offline non-linear edition software are unable to read. Thus, all videos need to be converted – luckily, there is a lot of free software nowadays that are able to do it! By the way, an observation here is pertinent: between August 2006 and March 2007, those free software were hard to find on the Internet and it was extremely difficult transform .flv in other codecs and vice-versa (From the beginning YouTube has been equipped with an auto-application to convert different video formats to its Flash structure); I am talking about this specific period – approximately seven months long – because of my own empirical and professional experience: the first date refers to a job done during the production of *visual jockey* (a live images remix performance) making use of YouTube videos, and the second one is related to another online audiovisual application launched at that time, "Manifest Yourself Online"[3], webcam-self part inside "Manifest Yourself [everybody is an artist] – Mobile WebTV Live Broadcast" project (2006-2008).[15]

Besides, this digital handicraft work tends to disorganization, i.e. the video downloaded loses its address (url) from YouTube and that link used to be a part of an online structure; without this source, there's no rigor inside this academic project, because in this area the paths are an important research part.

4.2. Learning, writing and quoting acts

Lawrence Lessig, Stanford professor and lawyer, presents wonderful arguments that help us understand Copyrights infringement in our "digital culture": while learning how to write, we are encouraged to read several authors, comprehend the meaning intended, quote them so that we are able to create our own composition; i.e. our talent in this media flows by utilization of good quotes to a new concept, we must do this construction from a pre-existing material comprehension we read during a research process.

The arguments are thus constructed based on a discussion about quotes and the intellectual circle to which this writing needs to be

[1] Paul D. Miller apud Manovich

[2] "For instance, Web users are invited to remix Madonna songs at http://madonna.acidplanet.com/default.asp?subsection=madonna." apud Manovich

[3] this online application – and independent platform – had cancelled its video-files storage and interactive functions during the second semester of 2008, because high costs of *Flash Communication/ Media Server*.

presented assesses the value of the study through its analyzed paths. The success of a text is closely related to the interaction between the quotes and the style of the text itself, i.e. the quotes selection displays the acquired knowledge from original text (source): "Only the most careful reader could construct from the text he read another text that explained it."[16]

In the world of the Visual Arts in the 1920's, in Europe, and in the 1960's, in the USA, the absorbing and re-appropriation of pre-existing material in society used to be called "collage" and "pop art" respectively. Manovich establishes an interesting difference between "collage" and "remix [electronic and software collage]":

> *"To use the terms of Roland Barthes, we can say that if modernist collage always involved a 'clash' of element, electronic and software collage also allows for 'blend.'* [13]

Let us examine some of Lessig's ideas about "remix" and audiovisual matter from the perspective of quoting:

> *"So here's the question I want you to focus on as we begin this chapter: Why is it "weird" to think that you need permission to quote? ... Why is an author annoyed (rather than honored) when a high school student calls to ask for permission to quote? ... what happens when writing with film (or music, or images, or every other form of 'professional speech' from the twentieth century) becomes as democratic as writing with text? As Negativland's Don Joyce described to me, what happens when technology 'democratiz[es] the technique and the attitude and the method [of creating] in a way that we haven't known before. . . . [I]n terms of collage, [what happens when] anybody can now be an artist'? ... should the norms of "quote freely, with attribution" spread from text to music and film? ... we should think more carefully about why this right to quote— or as I will call it, to remix— is a critical expression of creative freedom that in a broad range of contexts, no free society should restrict. ... there are important reasons why we should limit the regulation of copyright in the contexts in which RW creativity is likely to fl ourish most. These reasons reflect more than the profi t of one, albeit important, industry; instead, they reflect upon a capacity for a generation to speak."* [16]

Lessig, a founding board member of *Creative Commons*, is one of the exponent names in a politicized "digital culture". This new culture, which originated from the so called digital media and new technologies, is invested with re-appropriation practice of cultural material in a creative environment and also aims at spreading information and developing knowledge. Lessig becomes an advocate of a reality which does not comply with the laws to which culture is subject to in a Digital Network Society: a reality that brings into being a huge amount of new multimedia thinkers (digital craftsmen), thanks to sharing and re-appropriation.

Although this reality is extremely complex when we consider the laws that regulate the entertainment and art industry, Lessig's take appeals to millions of young digital crafters, who make the act of expressing themselves in audiovisual form as trivial as speaking or writing.

4.3. The audiovisual writing(s)

When Lessig advocates a new legislation concerning these audiovisual re-appropriations in our "digital culture", he somehow returns to a cinematographic discourse previously existent in Cinema Theory:

> *If the movie-making mechanisms were not widely available to everyone, as Metz asserted in the 60's: "All English speakers as of a certain age master the usage of the language and are therefore capable of producing sentences, but the ability to produce a film discourse is dependent on exposure and education". Maybe he was anticipating a Digital Network Culture when he says that "We can foresee a future society in which all citizens will have access to a code of cinematographic production", where we can conceive the "cinematographic" as the audiovisual of current new media.* [17]

The verbal writing has – as it happens in every academic research – a reference and social learning promoted by readings, quotes and re-writings; thus, why must audiovisual creation follow just a marked logic? Lessig enters the 21st century – 40 years after *Détournement*-production by Debord – arguing in favor of audiovisual material re-utilization, as opposed to an old *Copyright* model, in defense of a change in laws which are not attuned to this already transformed and widely distributed culture, mainly on the Internet. Let us imagine then a sociable-cultural reality where any person is able to produce his own discourse (rhetoric) not just in a verbal writing manner but also audiovisual and in a multimedia system.

Détournement is a French word that means both "swerve" and "steal". It is a key-word used by the Situationist Group, of which Guy Debord was the main exponent. This movement proudly explored the dichotomy in this word to elaborate their "revolutionary practices".

> *"In détournement, an artist re-uses elements of well-known media to create a new work with a different message, often one opposed to the original. ... Détournement is similar to satirical parody, but employs more direct re-utilzation or faithful mimicry of the original works rather than constructing a new work which merely alludes strongly to the original. It may be contrasted with recuperation, in which originally subversive works and ideas are themselves appropriated by mainstream media. ... Détournement's use by Barbara Kruger made many people familiarwith the technique, and was extensively and effectively used as part of the early HIV/AIDS activism of the late 1980s and early 1990s. Examples of contemporary detournement include Adbusters' "subvertisements" and other instances of culture jamming [eg. when quotations by the former United States President George W. Bush are combined with haiku-like phrases to produce a larger work intended to subvert the original source.]"* [18]

Hence, Guy Debord is cited in this short article because he made two films strictly based on pre-existing film material re-appropriation, like he explained in a short text in 1989 called ***The Use of Stolen Films*** regarding "Society of Spectacle" movie (1973):

> *"(I'm talking here primarily about the films that interrupt and punctuate with their own dialogues the text of the spoken "commentary" derived from the book) ... In A User's Guide to Détournement (Lèvres Nues #8) we already noted that 'It is thus necessary to conceive of a parodic-serious stage in which detourned elements are combined . . . in order to create a certain sublimity.' Détournement is not an enemy of art. ... Thus, in the film The Society of the Spectacle the (fiction) films detourned by me are not used as critical illustrations of*

an art of spectacular society (in contrast to the documentaries and news footage, for example). On the contrary, these stolen fiction films, external to my film but brought into it, are used, regardless of whatever their original meaning may have been, to represent the rectification of the 'artistic inversion of life.' The spectacle has deported real life behind the screen. I have tried to 'expropriate the expropriators.' [Here Debord explains in a few words his views on the re-appropriation of material inside his cinematographic work] ... The Rio Grande [for example] sequence is intended to evoke historical action and reflection in general. Mr. Arkadin [Orson Welles movie] is at first brought in to evoke Poland, but then hints at authentic life, life as it should be. The Russian films also in a sense evoke revolution. The American films on the Civil War and Custer are intended to evoke all the class struggles of the nineteenth century; and even their potential future." [19]

After forty years of Debord's and several audiovisual artists[4] gambling in found footage material re-appropriation to create a "new culture", this issue comes up even stronger, because nowadays this kind of practice is still being labelled as "piracy", i.e. it points out to a symbolical pressing necessity of this situation supported by young handcrafters within a "digital culture".

"it is very possible that the remix culture, which right now [2007] appears to be so firmly in place that it can't be challenged by any other cultural logic, will morph into something else sooner than we think." [13]

Although remix (and YouTube) is not anymore an innovation, the discussion about its implications is still up to date and a new generation hopes to live free of infringement situations.

Therefore, if academic thesis and essays texts are written bearing in mind quoting and bibliographic references, my project assumes that there are other alternative cognate scientific forms in our YouTube era; i.e. this proposal aims at an audiovisual meta-language as a methodological system for noting, recording and remixing cultural codes available on the Internet. After all, in those digital-culture virtual spaces, method is as important as the chosen media.

5. YouTubteca, YouTubeMix, YouToRemix, RemixPress

Everyday an enormous amount of information is posted on the web, but how can it be organized? Several online applications already take care of these ever increasing paths that sometimes stray or converge, but always link to one another. How can we make any audiovisual quote (link) easier inside a remix process? Maybe the expression should not be "make it easier", but rather "organize", assign a path for what has been researched for analysis and further creation. How can we stimulate people to create information from such varied paths in an environment where there is also the possibility of remixing this material and making it accessible to other users?

The application proposed here – the tech-methodological part of my master degree research –, is inserted in the so-called "digital culture" and represents a technological innovation, since it introduces an online audiovisual practice concept. As a study, it shows how videographic production and creation can be achieved using online videos. It may also be used as an important tool in learning environments, helping create rhetoric and expressiveness while working with both content and aesthetic aspects. The usage of this online application aims at promoting research and motivating networked participatory, shared experiences to develop cultural and digital skills.

Thus, I consider YouTube – and other current online audiovisual platforms – as a tool that opens a whole new spectrum of artistic and educational possibilities rather than just a video storage device. After all, the secret to make the student – or user/spectator – actively engaged in the process of creation is to transform him in a co-author .

"These paths [which facilitate movement of information between people] stimulate people to draw information from all kinds of sources into their own space, remix and make it available to others, as well as to collaborate or at least play on a common information platform. Barb Dybwad introduces a nice term 'collaborative remixability' to talk about this process: 'I think the most interesting aspects of Web 2.0 are new tools that explore the continuum between the personal and the social, and ***tools that are endowed with a certain flexibility and modularity which enables collaborative remixability — a transformative process in which the information and media we've organized and shared can be recombined and built on to create new forms, concepts, ideas, mashups and services."*** [20](my highlights)

Let us refer to Manovich once again, "helping cultural bits move around more easily"; the proposed application, whose primary function was to facilitate the production and creation of the *Audiovisual Rhetoric* and its source noting, is inserted in a data terminology linked to "web 2.0":

"The Web of documents has morphed into a Web of data. We are no longer just looking to the same old sources for information. Now we're looking to ***a new set of tools to aggregate and remix microcontent in new and useful ways."*** [20] (my highlights)

The principle of application usability is based on three current online audiovisual tools. While two of them use metadata, i.e. they are constructed having YouTube videos as a starting point (via url), the third one proposes a real time audiovisual creation controlled by the connected user (it's an online live remix possibility with sounds and videos that are previously stored in its server plus real time images captured from user's webcam). The references listed below present the development path of my application:

- **YouCube** [21]
- **YoooooouTube** [22]
- **Vjiar: Web-Vj'ing-Cam** [23]

This new online audiovisual live remix application project aims at being a data organizer (a covered audiovisual path) for its own

[4] At the same time, Carlos Adriano, an audiovisual producer who presented his doctorate thesis at USP (2008), mentions other artists related to *Found Footage* practices as well as its main usages: analytical, structural and critical(in which we can include Debord and Godard). For a deeper understanding of this taxonomy see Brenez and Chodorov (2002). For a more detailed analysis of the practices of *found footage in the first half of the 20th century* see Jay Leyda (*Films Beget Films*, 1964).

online switcher to create live ***Audiovisual Rhetoric*** ; i.e. the intention is to establish an online method of audiovisual writing with any digital videos (data) stored at YouTube: a methodological elaboration for online rhetoric such as quick audiovisual quotes, like postings in a blog. The application may be a kind of audiovisual blog, where writing (text) is created through the remixing of videos linked in YouTube.

After the user is connected through his login, he gets access to the application, which is divided in three stages: 1. Space to organize data (where he can type his video links: url addresses selected from YouTube); 2. User will be able to select a specific footage inside YouTube videos (the possibility to control time position of a video clip to be played and remixed), these footages can be put in loop, reverse and have their playing speed changed (faster/slower feature); 3. This stage is the most awaited one: it's the remix environment per se, designed in a *patch* concept of creation which enables the user to assign dynamic connections between video clips and app functions (see the following several functions).

Print screen of "Vjiar: Web-Vj'ing-Cam"[5],
during interactivity process (live video remix)

Therefore, it's based on a vj-style setup which will allow non-professionals to mix, apply effects to video clips and even scratch, everything done online and in real time. Application allows the adding of different functions to the video stream, an effect – for example – can have up to X parameters. These parameter values are set by keyboard, automated in different ways.

5.1. *Referencial functions:*

- User selects YouTube videos and links them (via url, like tested at *YouCube* and *YoooooouTube*);
- User can play, trigger and remix those video clips and sounds via keyboard. He can apply a series of audiovisal effects based on analogical video switchers and also present at some digital visual jockey software, like colorize (RGB) and opacity (both with lowpass/highpass filter), image repeating (non-controllable parameters of number of columns neither of rows), luma-chroma key etc So, user creates a new composition in real time and he can play also using his own image from webcam (like experienced at *Vjiar: Web-Vj'ing-Cam*);
- It is important to have a "record function"[6], i.e. it records this live composition to online server as a new video clip; it's done without interrupting live experience. These new video clips are created in a user space (own channel inside application) where he can choose how to publish it: draft, private (with a password) or public, like it occurs in a blog;

5.2. *New functions:*

- the possibility to control time position of a video clip;
- lettering-text input: user can write moving texts over the images;
- interface design: possibility to customize and personalize the own channel (colors, background image, title, text etc);
- footage can be put in loop, reverse and there is also a speed control.

5.3. *Future functions:*

- SCRATCH: video frames backward and forward;
- MIDI CONTROLLER: for any device to control the app; user will map preloaded video clips to be enabled trigger and knob;
- DOWNLOAD: possibility to storage offline the new compositions, live creations;
- Upload a remixed video straight into your own channel at YouTube.

Finally, this new app intends to be a simple online platform-system based on metadata technology in which some of the characteristics found in several current vj'ing software are brought together. Its target, primarily, is to create an online *Audiovisual Rhetoric*.

As this is an open source project, we hope that the online community comes together to develop new tools and functions according to needs and eventually share it again. (in the same lines of the *wordpress* project).

6. ACKNOWLEDGMENTS

I'd like to thank my superviser Esther Império Hamburger and the funding institution CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) which have been provided support for the development of my master degree at Art & Communication School in University of São Paulo (ECA-USP).

[5] Demo version presented at Transmediale Festival site (2006). Keyboard is used to apply effects and trigger images and sounds

[6] According to what happened at "Manifest Yourself Online" (@Manifesto21/2007)

7. REFERENCES

[1] http://en.wikipedia.org/wiki/WordPress

[2] http://wordpress.org/

[3] http://www.videonuze.com/blogs/?2010-02-09/comScore-Data-Shows-2009-Was-a-Blistering-Year-for-Online-Video-Slides-Available-/&id=2425

[4] http://www.comscore.com/Press_Events/Press_Releases/2006/10/YouTube_Worldwide_Video_Streams/%28language%29/eng-US

[5] "Manifeste-se [todo mundo artista] – Mobile webTV Live Broadcast"<http://www.manifesto21.com.br/manifestese2006/projeto.htm>

[6] http://news.cnet.com/8301-10784_3-6133076-7.html

[7] http://techcrunch.com/2006/10/09/google-has-acquired-youtube/

[8] JENKINS, Henry. 2009. O que aconteceu antes do YouTube? in YouTube: digital media and society series. (Burgess & Green org.) [brazilian version] Aleph : São Paulo.

[9] *Vídeo Gratias.*2007. in Cadernos Videobrasil, número 03

[10] http://www.videonuze.com/blogs/?2010-02-09/comScore-Data-Shows-2009-Was-a-Blistering-Year-for-Online-Video-Slides-Available-/&id=2425

[11] HARTLEY, John. 2009. Utilidades do YouTube: a alfabetização digital e a expansão do conhecimento in YouTube: digital media and society series. (Burgess & Green org.) [brazilian version] Aleph : São Paulo.

[12] http://www.youtube.com/t/howto_copyright

[13] MANOVICH, Lev. 2007. What comes after Remix? <www.manovich.net/DOCS/remix_2007_2.doc>

[14] http://remixtheory.net/?page_id=3

[15] http://blog.manifesto21.com.br/2007/11/14/mobile-webtv-live-broadcast-txt-english2/

[16] LESSIG, Lawrence. 2009. REMIX – Making Art and Commerce Thrive in the Hybrid Economy. Bloomsbury Publishing PLC : London

[17] SZAFIR, Milena. 2009. Master Degree 1st Report, May 2009. Communication and Arts School in University of São Paulo.

[18] http://encyclopedia.thefreedictionary.com/detournement. (accessed in 2010)

[19] DEBORD, Guy. 1989. The Use of Stolen Films <http://www.bopsecrets.org/SI/debord.films/stolen-films.htm>

[20] MANOVICH, Lev. 2005. Remixability <www.manovich.net/DOCS/Remix_modular.doc>

[21] http://www.universaloscillation.com/youcube/?3985

[22] http://www.yooouuutuuube.com/v/?rows=24&cols=16&id=nxqt1rpj-bU&mode=shuffle&startZoom=1

[23] http://www.manifesto21.com.br/game

Mobile TV: Cultural Based Design Evaluation

Menandros Apostolidis
Oxford Brookes University
1st line of address
2nd line of address
Telephone number, incl. country code
chriskel@otenet.gr

Dr Maaruf Ali
School of Technology
Oxford Brookes University
Oxfordshire, OX33 1HX, UK
+44 1865 48 3446
mali@brookes.ac.uk

ABSTRACT

This paper examines and presents the finding of the correlation between culture, education and Mobile TV's design principles. Design guidelines are formulated by pinpointing the limitations of handheld devices that concern viewer's experience and combining them with an investigation on how culture and education is affecting the human perception of interfaces and interactivity. Additionally, the guidelines are incorporated into a video player in order to test and evaluate them. The player was distributed and with the help of a questionnaire a survey was conducted. The results of the survey suggest that the player was successful in giving a pleasant experience. The results also reveal a correlation between culture and Mobile TV software design. Some differences in preference according to the education level were found. Finally, the findings are discussed in the context of future research and the prospects of Mobile TV with further suggestions offered.

Categories and Subject Descriptors

H.5 [INFORMATION INTERFACES AND PRESENTATION]: User Interfaces - Evaluation/methodology, Prototyping, user-centred design.

General Terms

Design, Experimentation, Human Factors

Keywords

Mobile TV, User-based design, Evaluation

1. INTRODUCTION

The subject of this study is an investigation of Mobile TV HCI (Human Computer Interaction), specifically, the examination of the effect of certain user characteristics in Mobile TV's interaction design. Previous research regarding this from a user based prospective was not extensive [1]. As a result, this new research was conducted from a user centred standpoint that attempts to take into consideration the factors of culture and education. According to Love, the user characteristics that may affect the process of use are some of the elements that play an important role for designers to understand the correlation among humans and mobile devices [2]. Initially, culture influences the way people perceive the surrounding environment and respond to it with a different style [1]. In addition, studies have shown that there is a correlation between culture and mobile interfaces [3]. Respectively, education is also a factor that affects interaction in Mobiles according to other research [4].

Hence, the purpose of the study is to make an effort to provide design engineers with specific results in the form of guides. These design suggestions can contribute to the unproblematic construction of Mobile TV software that will provide a high level of Quality of Experience (QoE) and be commercially successful. Furthermore, by examining the technologies used in Mobile TV and how they affect the audience's experience an attempt was made to correlate the two investigations. As an emanation of this research, design guidelines were formulated. The proposed guidelines were incorporated into the features of an Adobe Flash® based prototype constructed for handheld devices with specific system requirements. The capability of achieving an Internet connection and a QVGA screen resolution are among those requirements. The type of the application is an on-demand video player that delivers streamed video clips from a supposed provider. Finally, the video player was distributed to a group of people. The user sample tested the software and with the use of a questionnaire results were formulated and analyzed.

2. METHODOLOGY

2.1 Literature Research

Mobile devices are a part of a fast evolving industry with a highly developed technology behind it. Mobile TV technologies were examined and several findings were discovered. There are many limitations that constrain the user experience. Device performance, power consumption, signal reception, screen size, interoperability in input methods and technical specifications are some of the factors that result in a poor QoE [5][6].

The next part of the research was to examine the human dimension of Mobile HCI. Mobile device users have many idiosyncrasies. Mobile TV can be watched anywhere, at any time and in any context with a user that is always on the move and who can be easily distracted. The investigation on personal characteristics was mainly concentrated on the factors of culture and education. The Mobile TV market is one of global scale. Most of the software is being created in the USA and thus the designers use metaphors, representations, colours, navigation logic and text layouts all adjusted to the American perception [3]. Research that

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
EuroITV2010, June 9th-11th 2010, Tampere, FINLAND.
Copyright 2010 ACM 978-1-60558-831-5 /10/06…$10.00.

has been conducted concludes that Westerners can focus on a single object, recognize its properties and categorize it (*analytic perception*) while Easterners have the ability to detect more background information and detect relationships between objects (*holistic*)[7]. Finally, a case study that considered these differences was conducted on 20 subjects (10 Westerners and 10 Easterners) and showed that there is indeed a correlation between culture and icon recognition in mobiles [3]. Therefore, cultural sensitivities must be addressed during the software design process.

Analogous studies that concluded in specific design rules concerned other forms of interactive TV (iTV) and handheld devices such as the ones conducted by Chorianopoulos [8] and Gong [9] respectively. The analysis of all of the above resulted in the following guidelines:

1. Provide interactive features
2. Allow personalisation
3. Provide with validation feedback
4. Allow users to have control
5. Create a consistent interface theme and navigation
6. Rely on connectivity as less as possible
7. Tolerance in errors
8. Design for users with short attention span
9. Design for small devices
10. Design for spontaneous interruption and resume functionality
11. Provide a high level of hierarchical information and allow the user to choose
12. Provide an untraditional content delivery

2.2 Prototype

After establishing the main rules of interface design and interaction, the next stage of the research was to apply those rules to an application prototype. The application has a five-button menu that signifies the basic actions as shown in Figure 1 and 2. The functions include: a "watch-now" option that instantly plays a random clip; the ability to see the available clips categorised; the option to search for a clip and customise the application. The customisation options include: language selection, theme selection and categories filtering. Moreover, the clips have full playback control. The user can: play, pause, stop and hop the clips. Pausing serves as a halt of the entire application. To adjust the sound, the viewer can use a volume slider alternately to his device slider. Also, an information button is positioned next to the thumbnail of every clip and when pressed the application opens an Internet browser window and redirects the user to a relevant webpage. The graphical interface is consistent and it's based on a grid layout. The icons are designed to have comprehensive metaphors and the first theme (Figure 1) has buttons with a rollover text while on the second theme the rollover text is absent. Moreover, a small vibration of the device each time a command was given by the user was intended to give a validation feedback.

These synoptically presented prototype features were chosen according to the previously mentioned guidelines. However, not all of the initially design goals were achieved due to problems encountered during the development process.

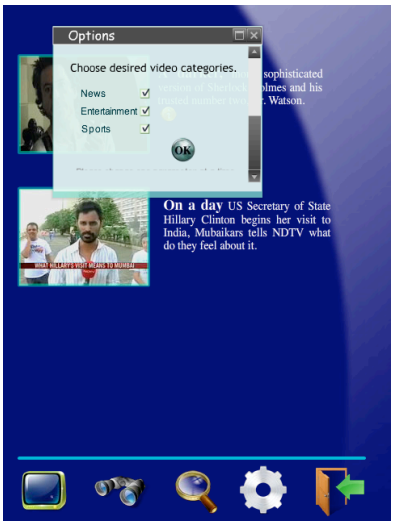


Figure 1. The prototype showing the settings panel open.

Figure 2. The prototype during video playback with an alternate visual theme loaded.

2.3 Evaluation

The next stage of the study was to distribute the application to a number of users in order to evaluate it. An amalgamation of sampling techniques was used to select the users. Firstly, the sample is biased. It was imperative to have people from both Western and Asian cultures, males and female, users that can understand English or Greek and from various ages and educational backgrounds. In that way features for specific user groups could be tested and the external validity would be improved as the conclusions could be generalized for a larger population. Furthermore, stratified (Table 1) sampling method was also used. The number of users was divided into strata in order to form the population into groups and sample from each group. In that way the final user group is more representative and reflects the target population. The strata are formed according to criteria that are of interest to the case study, culture and education. Opportunity sampling was also a part of the amalgamation as the

users were also selected because they were easily available. The size of the sample was ten users. It is relatively small to generalize conclusions, thirty would be a more accepted sample size. However, there are researchers who believe that five to ten participants are a sufficient number to get the main results [10].

Table 1. The eight strata used to perform the sampling.

Greek, Highly Educated	2
Greek, up to Further Education	2
Indian, HE	1
Indian, up to FE	1
British, HE	1
British, up to FE	1
Chinese, HE	1
Chinese, up to FE	1

In order to achieve final results with internal and external validity the users were carefully selected so that they were not biased or extremely friendly with the service.

The survey was performed with the use of a questionnaire (Likert questionnaire) and a short informal interview with the form of a discussion followed. The questionnaire consisted of sixteen questions as illustrated in Table 2. The users tested the application in their own time and location and filled the form after a few minutes of use. No guidance was provided to the participants on how to use the application. However, they were informed that they could ask for assistance if they would experience any problems. When they returned the filled form a short discussion took place. The participants were asked about their overall experience, if they found any significant faults and what would they add or do differently.

With the scaler questionnaire the users were capable in declaring the extent on which they agree or disagree with a statement. However, the majority of the answers were converted to a binary yes or no response. This was chosen because it is more convenient for the data analysis to have fixed answers and specific replies of approval or disapproval for each aspect of the video player. The sample size is not sufficient to create statistics for each degree of approval. Thus, the positive answers were converted to a "yes" and the negative ones to a "no".

Table 2. The sixteen questions used in the questionnaire.

1	How old are you?
2	What is your sex?
3	In which country have you grown up?
4	What is the level of your education?
5	Do you use mobile TV services?
6	Did you enjoy the specific video player?
7	Was it similar to normal TV watching?
8	Was it easy to use?
9	Did you stumble across any problems?
10	Would you add any additional features to it?
11	Do you prefer text or icons in interfaces?
12	Did you like the options of changing the looks of the app?
13	Do you think that the option of filtering the categories was useful?
14	Would you pay for a service with those characteristics?
15	Would you prefer to watch live TV?
16	Where would you watch it?

3. RESULTS

After gathering the result and analyzing them with the help of statistical software the following results were found. Fifty per cent of the participants used Mobile TV services in the past. Additionally, 70% declared that their experience was not similar to traditional television watching while 40% would actually prefer the second. Moreover, 90% stated that they enjoyed the player in overall and that it was easy to use. However, 30% of the users did encounter some type of problem. All of the users thought positive of the personalisation options although 30% did not agree with the categories filtering feature.

As Figure 3 illustrates, participants who grew up in Western countries have the tendency of preferring text over icons, or the combination of the two, whilst Asian participants have the opposite tendencies. This can be explained from the fact that the first icon menu set in the application has a more concrete design with more obvious icon metaphors and rollover text.

Correlating the culture factor with the user preferences on the customisation features lead to the observation that mainly Asians approved the categories filtering feature while only half of the Westerners did so. The reason of this can be that the three categories (music, news, sports) are not so famous among the Asian users. In addition, all the participants were satisfied with the theme change option.

Furthermore, the education parameter when correlated with the same answers did not form very interesting results. Users with a higher education approved the categories filtering feature slightly more. The unexpected result is that in the icon or text preference question, the users with a lower education level tends to prefer text in interfaces slightly more.

Another observation includes the fact that almost everybody stated that they were satisfied with the application although only four would purchase it. Additionally, two users that answered positive about the service purchase also answered that they would prefer to watch live television instead although they would use the mobile service at home.

Finally, the informal interview also provided with some interesting information. The participants suggested additional features for the application. Western culture users focused on practical issues, watching news bulletin for example after a notification and adding more categories of videos. Another suggestion concerned an omission of the prototype; the language of the settings panel would not change according to the application's language. Other suggestions included the addition of a home button and rollover text for the rest of the buttons besides the main menu. Finally, users stated that the application was easy to use while on the go and additional features could help although, as they said, they would not like it complicated.

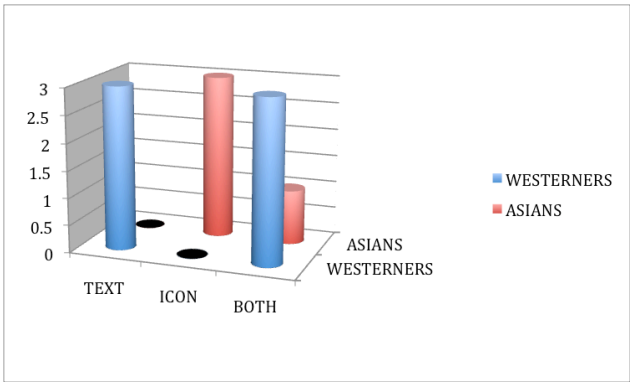


Figure3. Preference on icons and text by culture.

4. DISCUSSION AND FURTHER SUGGESTIONS.

The survey indicated that overall the navigation and interface was pleasant for the test users and the general structure of the prototype was acceptable. Having analyzed the user feedback conclusions can be drawn about the validation of the design rules. Approval of participants in their majority concerned the interactive features and they commented that there was a good balance between features and mobility. A feature that allowed the application to pause was also high in the approval rating thus confirming that the "Design for spontaneous interruption and resume" guideline improved the viewer's experience. "Provision of video with untraditional methods", is a guideline that was not validated. The prototype was designed to combine a television "feel" in a video carousel application. Most participants answered that they would prefer to watch traditional live TV. This finding may indicate that users are not very receptive in accepting new ways of video delivery.

Customisation and personalisation have proven to be an essential design choice in order to address cultural sensitivities. However, the Chinese participants encountered some difficulties - further research is needed to explain the cause of this outcome. In contrast with the cultural factor, no significant conclusions can be drawn when correlating the education level of the participants and their answers. Additional research is required that will focus only on this parameter with a bigger and more representative sample. The sample selection may have influenced this finding. It is also a fact that almost all the users are experienced with mobiles nowadays (in Asia and the Western world). Moreover, less than half of the participants would purchase the service. If we consider some of the comments of the participants during the interview, "the application reminded them of YouTube" another conclusion can be drawn from these two facts. The participants would not purchase a service with the same characteristics that a home computer or laptop could offer.

However, additional research is required in order to specify all the elements that are correlated with cultural-related perception; colours, icon metaphors, information layout and sounds all play a significant role in user-centred design. A future study should also investigate right-to-left writing cultures.

ACKNOWLEDGMENTS
A warm thank you to Konstantinos Chorianopoulos for his advice and contribution.

4. REFERENCES

[1] Kim H. J. and Lee K. P., "Cultural Difference and Mobile Phone Interface Design: Icon Recognition According to Level of Abstraction". Mobile HCI'05, 19-22 September, 2005, Salzburg, Austria. Paper available from: http://dpl.kaist.ac.kr/web_wiki/images/9/99/Kimj2005.pdf [accessed on Feb., 2010].

[2] Love, S., Understanding Mobile Human-Computer Interaction: A Psychological Perspective. Buttworth-Heinemann, 2005, ISBN 978-0750663526.

[3] Li H., Xianghong S, and Kan Z., "Culture-Centered Design: Cultural Factors in Interface Usability and Usability Tests", 2007. Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing 2007, SNPD 2007, Vol.3, p.1084-1088, ISBN 978-0-7695-2909-7.

[4] Medhi I., Sagar, A. and Toyama, A., "Text-Free User Interfaces for Illiterate and Semi-Literate Users". Information Technologies and International Development, Vol. 4, Issue 1, p.37-50, MIT Press, 2007. ISSN 1544-7529. Paper available from: http://research.microsoft.com/en-us/people/indranim/text-freeui.pdf [accessed on Feb., 2010]

[5] Chorianopoulos K., "Personalized and mobile digital TV applications", Multimedia Tools and Applications (2008), Vol.36, Nos.1-2, Jan 2008, p.1-10, published online: 27 January 2007, Springer Science Business Media, LLC 2007. ISSN (online): 1573-7721, ISSN (print): 1380-7501.

[6] Knoche H. & Sasse M.A., "Getting the Big Picture on Small Screens: Quality of Experience in Mobile TV", Interactive Digital Television: Technologies and Applications, eds. G. Lekakos, K Chorianopoulos and G. Doukidis, 1st ed. London: IGI Publishing. Chapter XIV: p.242-260, 2008.

[7] Nisbett, R.E. The Geography of Thought: How Asians and Westerners Think Differently…and Why. The Free Press. New York, 2003. ISBN 978-0743216463.

[8] Chorianopoulos, Konstantinos, "User Interface Design Principles for Interactive Television Applications", Int. Journal of Human-Computer Interaction, Vol.24, No.6, 2008, p.556-573. Paper available from: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=47D96812D4F588EE7DB49BAE8F40549A?doi=10.1.1.60.1291&rep=rep1&type=pdf [accessed on February, 2010].

[9] Gong J. and Tarasewich P., "Guidelines for Handheld Device Interface Design", Proc. of the 2004 DSI Annual General Meeting, 2004, p.3751-3756. Paper available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.87.5230&rep=rep1&type=pdf [accessed on February 2010].

[10] Neilson J., Usability Engineering (Interactive Technologies), published by Morgan Kaufmann, 1st Edition, 1993, ISBN-13: 978-0125184069.

WORKSHOP 1

Interactive Digital TV in Emergence Countries (BRIC)

Editors:

Prof. Vicente Ferreira de Lucena Jr. (University of Amazonas, Brazil)
Prof. Lin Gui (Shanghai Jiaotong University, China)
Prof. Rajendra Akerkar (Technomathematics Research Foundation, India)
Dr. Drago Torkar (Jozef Stefan Institute, Slovenia)

Development of an Academic Laboratory of Interactive Digital TV in Amazon, Objectives and Main Projects

Vicente Ferreira de Lucena Jr
University of Amazonas -- Ceteli
Campus Universitário
Manaus – AM, 69077-000, Brazil
+55 92 3305 4680
vicente@ufam.edu.br

Waldir Sabino da Silva Jr
University of Amazonas -- Ceteli
Campus Universitário
Manaus – AM, 69077-000, Brazil
+55 92 3305 4678
waldirjr@ufam.edu.br

Celso Barbosa Carvalho
University of Amazonas -- ICET
Campus Universitário "Moisés Benarrós Israel"
Itacoatiara – AM, 69100-000, Brazil
+55 92 3521 3519
celso@gta.ufrj.br

José Pinheiro de Queiroz Neto
Federal Institute of Education, Science & Technology
IFAM – Campus Manaus Distrito Industrial
Manaus – AM, 69075-351, Brazil
+55 92 3613 3530
pinheiro@ifam.edu.br

ABSTRACT
As the main goal of this workshop is to bring together researchers, educators, and industry related people working in the deployment of Interactive Digital TV Systems and Applications in developing countries, we decided to write about the establishment of a laboratory in the most important university of Amazon in Brazil that works with this technology. This laboratory is located in the CETELI/UFAM and was built to help the local technical community to make experimentation on Digital TV related themes. In fact, the most important goal of this technological center is the teaching of undergraduate and master of science students as it servers the electrical and computer engineering courses as well as the master of engineering program of the Federal University of Amazon. In the next pages a description of the main projects running in this laboratory is going to be made in order to give an insight about its potential and objectives.

Categories and Subject Descriptors
B.0 [**Hardware**]: General; D.2.0 [**Software**]: Software Engineering: General; I.0.m [**Computing Methodologies**]: General; J.7 [**Computer Applications**]: Computers in Other Systems;

General Terms
Management, Documentation, Performance, Design, Experimentation, Human Factors, Standardization, Languages, Theory, Legal Aspects, Verification.

Keywords
Interactive Digital TV Systems, Digital TV Standards, ISDB System, New Applications for Interactive Digital TV.

1. INTRODUCTION
The city of Manaus, located in the center of the Amazon Forest in Brazil, is home to the largest global free trade zone for the production of electronics and the manufacturing of digital television sets. The eighth largest city in Brazil, and the largest industrial park of its kind in South America, Manaus is a microcosm for the political, economical, and social transformations implicated in recent debates over Digital TV [1, 2]. Because of its economical characteristics, Manaus has several universities, technical colleges and institutes of research and development, and one of its challenges is the improvement of professionals and technologies to support the industrial pole. Indeed, the academy has many difficulties to follow the fast changes in technology and to attend the industries' necessities.

With the deployment of Digital Television in Brazil there is a growing demand for new services and contents to be provided, as well as many ways of integrating these new features with existing technologies [3]. The popularity of TV in Brazil opens a window of opportunities for many people working with related topics. New solutions and new business approaches are supposed to be born from this scenario, creating many new job opportunities [4, 5]. The University of Amazon, particularly the CETELI in Manaus, hosts a modern laboratory working with topics related to these new technologies. The Interactive Digital TV laboratory was built with governmental funding and is sponsored by the local industries, having a relevant importance in the teaching of new engineers and scientists.

In this report we are going to describe the most recent projects under development at UFAM. Indeed, application like T-Health, T-Government, T-Commerce, and T-Learning are some examples of value-added services (VAS) investigated in this laboratory. We start with the integration system between the Brazilian DTV model and the ODGi framework that lead us to implement some automation home features. The second project is an extension of this first and deals with health monitoring using the TV set as a comfortable user interface. Education and social insertion is a very important topic in Brazil, section 4 describes a T-Education project that wants to help game developers in their jobs. Finally, in section 5 a software engineering experiment focused on Digital TV is presented. With these descriptions we hope to give a brief overview of the themes under development in our institution.

2. COMMUNICATION SOFTWARE BETWEEN THE BRAZILIAN iDTV AND RESIDENTIAL DEVICES

The characteristics of new interactive Digital TV (iDTV) systems are assuming a very important role in modern life. In fact, these features are being increasingly expanded, from simple signal decoders to sophisticated devices that allow the execution of interactive applications related to the content displayed, providing services of all kinds like internet access, TV-Banking, TV-Mail, TV-Commerce, Games and so on [6, 7]. This new world is about to be established in the Brazilian market.

In a Residential Automation System, various consumer devices such as microwave, air conditioning, and temperature sensors, can be integrated into a single interface with which the user can control/monitor through operations available by them. Due to a wide range of communication technologies there are different ways for interconnecting them in a Home Network [8].

Nowadays the common way of expanding the functionalities of a Set-Top Box (STB) is to connect it with other devices equipped with some computational power such as the electronic equipments of a Home Network (HN). This is possible through the use of the STB as the central processing unit of networked devices architectures, also known as Residential Gateways [9]. The exploration of iDTV interactivity to access devices lead to many usage scenarios providing a networked entertainment environment to the consumer, allowing for example, sending a message to a cell phone in the house, turn on/off the lamp or fan or managing the energy consumption of the residence.

We developed a collaboration model between the Brazilian iDTV system and Home Networks exploring particular features of the Ginga declarative and procedural environments (Ginga-NCL and Ginga-J). This brings new applications and scenarios to the Brazilian market, allowing the consumers to immerge themselves into an interconnected home, using Ginga's interactivity to manage devices and sharing information among them. The development was split into two phases: (1) choosing the communication technologies and (2) choosing the proper approach for managing of the electronic devices in a Home Network.

We focus on Wireless (WiFi, Bluetooth and ZigBee) network due to the following reasons: no use of cables to connect the electronic devices, easy integration of several types of equipments, the low energy consumption, and the small signal range between the connected devices.

From (2), the Open Services Gateway initiative (OSGi) specification was chosen because the services are managed locally by this framework, the devices identification is offered by a Java interface, the discovery of the services is done by searching in a Service Registry, and it supports communication with other specification for Home Networking, such as UPnP and Jini.

Finally, we created a platform constructed on a component-based model for the Ginga middleware and the OSGi framework. The general architecture of this model is shown in Figure 1. In this figure are Ginga-J, Ginga-NCL and OSGi components and the software bridge between them.

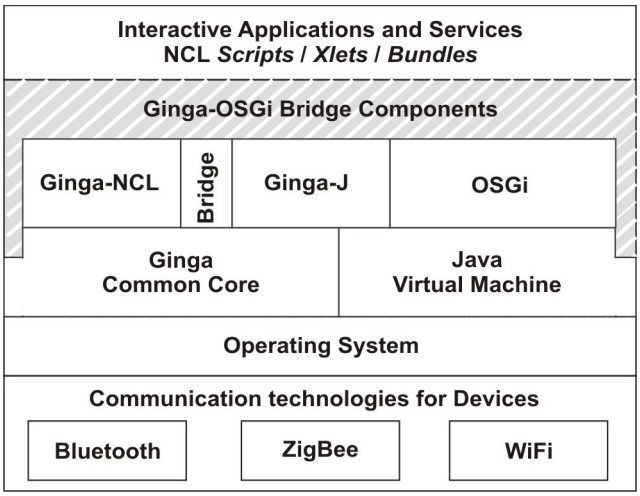


Figure 1. General Ginga-OSGi Architecture

Some functions for the new Ginga-OSGi model have been identified to meet the needs of integration between Home Networks and the Brazilian Digital TV. They are:

Allowing the registry and the discovery of OSGi services in a Home Network – Services are the basis of OSGi model. A service is an interface to operations provided by the networked devices. The service discovery in a Home Network is a key step for integrating with interactive Digital TV applications.

Interoperation Between OSGi services and Ginga functionalities – The environment must allow Ginga to access OSGi services. It is a difficult task because it requires an effort to integrate these different software platforms. Similarly, this model should enable the access of Ginga functions by OSGi.

Providing a useful description of the services to the user – Service description is the key to success in attracting a user. In addition to the information provided by OSGi services, such as name, version, and provider, other information should be released, such as device name and methods description.

We developed some strategies for collaboration between Ginga and OSGi: Ginga-NCL to OSGi, OSGi to Ginga-NCL, Ginga-J to OSGi and OSGi to Ginga-J. After the developing of Ginga-OSGi platform, we created three integration scenarios to validate this platform: (1) sending a message to TV screen by a mobile phone over Bluetooth; (2) monitoring temperature sensors located in some rooms of the house over ZigBee by iDTV application; and (3) managing a home access control by the Digital TV over WiFi.

The Ginga-OSGi model contributes to consolidate an iDTV-HN collaboration model, allowing the emergence of useful Home Automation applications and improving the user's experience and quality of life. In addition to the model based on the procedural specification (using Ginga-J middleware), the novelty of the work is in using declarative middleware features (through Ginga-NCL) to provide integration between script-based content and OSGi services. It considerably extends the scope of home automation applications that can be constructed using this new model.

3. HealthcareTV: AN INTEGRATION ARCHITECTURE AMONG MEDICAL DEVICES AND THE BRAZILIAN DTV

The architecture developed in this study used the specification of the Brazilian DTV middleware presented in [10], specifically the layer Ginga-J, which allows the Xlet Java applications to be executed.

The system collects data from a device equipped with a pulse oximeter, a temperature sensor and a blood pressure tester that sends information to a central computer at home equipped with a message alert system.

We can see in Figure 1 (B) this scenario based on an automated home, which has a local network integrated with an Internet router and a gateway to manage the devices in the house. Thus, taking advantage of this infrastructure, we insert the sensors and medical devices in the network using wireless networks connecting to the gateway.

The equipment is connected via Local Area Networking (LAN) and Wireless Local Area Networking (WLAN) and is used to transmit SMS messages via mobile phones integrated with the gateway. Completing the scenario, Figure 1 (A) represents the mobile system receiving warning messages via cellular network and the registration system accessing the data from home via WEB.

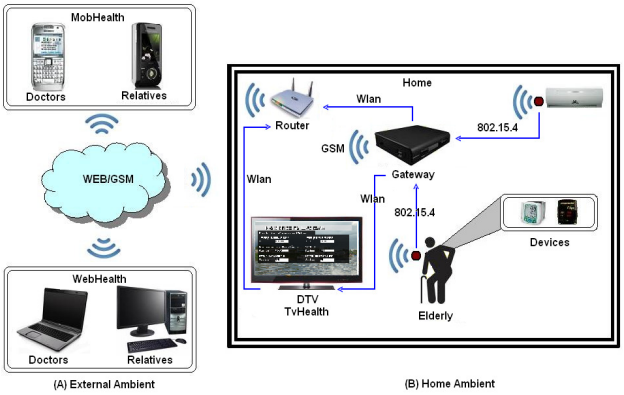


Figure 1: Scenario for a Monitored Residence

In our work we focus on the development of software architecture in X86 platform that encapsulates the virtual machine JAVA and JADE, installed in a Linux distribution, an application to monitor blood pressure, heartbeat and the patient's temperature. Moreover, our gateway has a WEB server, an application for digital TV and a mobile phone connected to provide SMS for doctors, nurses and relatives of patients being monitored

The system architecture has five layers, namely: Application, Decision, JAVA Virtual Machine (JVM), Operating System (OS) and Communication (COMM) (Fig. 2). The following description will show in details each layer of the system and their implemented functions.

The application layer is responsible for integrating all applications developed, such as WEB applications, mobile and DTV. This layer uses the JVM and WEB Server installed in the operating system, though the architecture which allows applications to be developed without the use of the JVM, for example particular applications developed in C running directly in the native operational system.

The decision layer is implemented using the JVM through Java Agent Development Framework (JADE) presented in [11] that is responsible for managing applications, exchange messages and updating the system status. This is done through the use of behavior supervision with pre-established agents which are capable of deciding whether to send SMS based on the changes observed in the sensors.

The decision layer is instructed to monitor the heartbeat received from the pulse oximeter, check the temperature received from the sensor and check the temperature and pressure measured by the patient through the blood pressure tester. If the parameters are normal, agents continue monitoring and do not send messages. However, if the parameters are abnormally exceeded, the supervisor agent sends a message to a second agent that triggers an SMS to mobile phone users and sends an alert to the Set Top Box which is presented on the DTV screen.

The protocol messages of agents have the format of the Agent Communication Language (ACL) defined in [12] and have a number of predefined fields such as: the sender of the message, the list of receivers, the communicative intention, the content and the ontology.

In the OS layer, there is a Linux operating system, customized for the operation of the gateway to healthcare, which has been configured in the JVM, WEB Apache Server, PHP, JADE, Zig Bee, and RXTX API for USB support, RS232 and SMSLIB API support for sending SMS. Figure 2 presents a high level division of the modules through all the layers presented.

In the application layer, the modules *DataRegister*, *HcMonitor*, *SendSMS* and *DTVAlarm* were implemented by fully sharing the same Extensible Markup Language (XML) file, located in the root directory of the gateway. This XML file is monitored by *HcMonitor* and classes *Oximeter*, *Pressure* and *Temperature*, responsible for collecting data from sensors.

Through the registration of users and sensors, the module generates the XML *DataRegister* collects initial information from patients and *HcMonitor* module does the monitoring and updating of such information. Figure 2 shows the architecture.

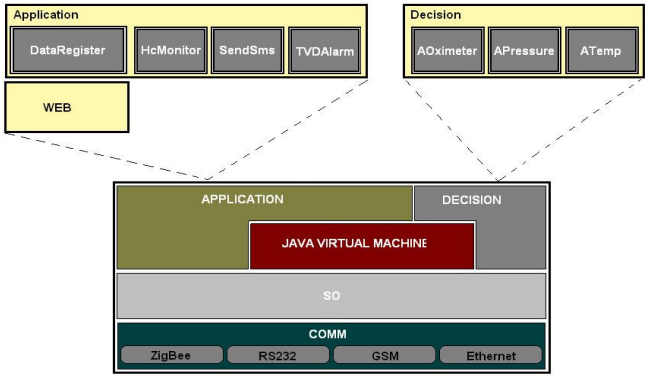


Figure 2: The General Architecture Dived by Layers

4. GameTV: A GAME FRAMEWORK FOR THE BRAZILIAN DIGITAL TV

The development of commercial interactive applications for Digital TV in Brazil is in its initial stages. Indeed there are a few consolidated experiences in creating educational applications for the Brazilian system available in the literature. Moreover, there are numerous difficulties in achieving quality education in large Brazilian regions like the Amazon State, where the use of Interactive Digital Television, that will be available to everyone in a few years, seems to be the most promising way for the educational inclusion of people who live in isolated, hard-to-reach cities. According to actual statistics more than 97% of the Brazilian families have access to the analog TV broadcasted signals and it is expected that at least the same amount of families will receive the digital signals in a very short time.

This scenario supports the proposal of this work, which is to investigate the environment of Digital TV and the necessaries characteristics for the development of educational games. The main goal is to propose a conceptual framework developed to help programmers to create games to digital television systems. It should assist the developer in the creation of these educational games, facilitating the construction of graphical user interface, handling of events and control of navigation.

As an experimental result we show that the platform contributes positively to the development of an educational game to teach engineering concepts, making the workflow of software developers and content television providers easier, abstracting the specific characteristics of the platform and the implementation of routines used in games. The easy usability of the framework permits that the teachers involved do not need to be professional software developers; they can use it to develop educational games to a specific content, including theory, goals and evaluation of the game itself, improving their classes and more interesting, reaching a large amount of families using the available Digital TV infrastructure.

4.1 Framework Architecture

The goal of this project was to propose a framework architecture for handling with games in Digital TV systems taking in account the Brazilian system and involving the study of the characteristics and requirements that such a framework should provide, within the limitations imposed by the Digital TV environment.

A study of the main characteristics of digital games was made in order to identify the features needed to be implemented and to propose an architecture for the conceptual framework. The requirements for the framework for digital games were defined as:

- It should provide support for the implementation of games running in Digital TV sets or STBs using the Brazilian system;
- The framework should use the paradigm of object oriented programming and give technical guidance to engineers and software developers in order to enjoy the benefits and ease of this paradigm;
- It should be constructed on basis of free software license, stimulating the research and share of ideas and experiences with the community;
- The Framework may initially operate in the Windows platform, but should be extended to support the operating system; GNU / Linux;
- The implementation of the Framework will be in Java because it offers portability, high performance and is object oriented.

The architecture suggested for the framework of games for Digital TV is explicit in Figure 1.

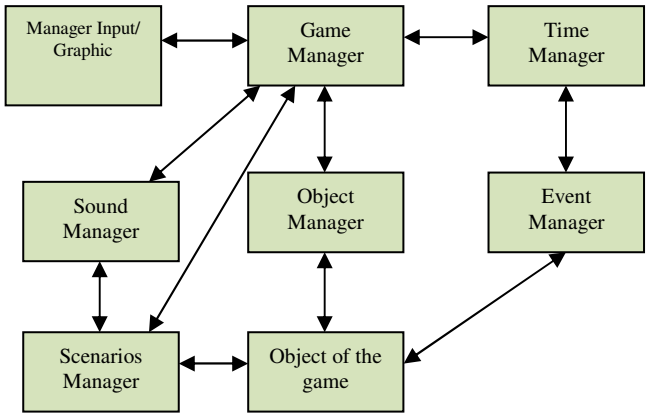


Figure 1. Conceptual model of the architecture of the Framework.

4.2 Implementation

Aiming to assist the development process of the architecture we present a survey of the main processes in relation to the functionality in the framework. Such processes have a simpler system functionality, and each process may involve one or more classes of control according to analysis. In the analysis phase we defined four basic processes:

- The main process is the system initialization and coordinator of other processes, responsible for creating the life cycle Xlet;
- Process of adding objects is the process responsible for creating objects such as text, image and sounds. Allowing the loading and manipulation of objects in memory;
- Procedure for handling events: responsible for all events in and out of the system by user interaction with the game;
- Process of converting the game to Digital TV: is the fundamental process as it allows the insertion of each object and performs the processing of them in Xlet to run the game in the STBs.

A tool named GameTv was developed and is responsible for the creation of graphical interfaces of games related to the main functional processes. The implementation is accomplished through the development of the following modules: Game Manager, Graphic Manager, Sound Manager, Event Manager, Object Manager, Scenarios Manager that will be presented in a future paper.

5. INTERACTIVE DIGITAL TV NAVIGATION SYSTEM PRODUCT LINE

An iDTV Navigation System combines program recommendation, sorting and retrieval to make it easier for the viewer to select programs based on various individual viewing habits [13], and is the guide to select services and applications, initiate interoperable applications, boot loading, and store user profiles [14]

In this Section we describe an iTV navigation system software product line (SPL), defining its reference architecture, and product backlog, that were described through a process called ScrumPL, that combines Software Product Line Engineering (SPLE) [15] and the agile method Scrum [16, 17].

This SPL is designed to create many products to be used in several parts of the word, for ISDB [18], DVB [19] and ATSC [20] standards and several languages, allowing a manufacturer to develop several iTV navigation systems at reasonable costs, enhanced quality and reduced time to market, each with similar user interface, the same recommendation system and Electronic Program Guide, and one of the available iTV standards (ISDB, DVB or ATSC), one language for a low-end, mid-end or high-end set-top-box market segment. Those reference architecture and product backlog are designed to be implemented and tested by several scrum teams.

5.1 ScrumPL Process

The ScrumPL process is composed of the Scrum lifecycle phases planning, staging, development and release [16]; and the SPLE processes domain engineering and application engineering [15]. During the planning phase, domain and application requirements are provided and added to the product backlog through the domain requirements engineering and application requirements engineering sub-processes respectively.

Those requirements are inputs to the staging phase, where the reference architecture is created and maintained through the domain design sub process, and the applications architectures are created and maintained through the application requirements engineering sub process. The architect, who is also the product owner, is responsible to create and maintain the reference and applications architectures, and to add each of its components, descriptions and interfaces to the product backlog as product backlog items.

The development phase starts when the product backlog is prioritized and made ready for estimation. The pre-game phase is finished and the Scrum team works according to Scrum process [17]. The results from Scrum team activities are potentially shippable components realized and tested, as well as unit and integration test cases (from domain tests and application tests) used to perform Scrum acceptance tests in this sprint (a 30 days iteration), and to be used in future regression tests. Those test cases are reusable.

During the 30 day sprint, the Scrum team can provide the domain realization goals: the detailed design and implementation of reusable software assets; or provide the application realization goals: applications that can be tested and brought to the market after ensuring sufficient quality.

Although Scrum states that the product owner is not able to change the product backlog items being realized by Scrum teams, the architect, as product owner, can make any changes in other product backlog items. Those changes are made due to, for instance, problem reports, new requirements, changes in interfaces, defects in interface descriptions, issues in realizing domain and application artifacts, and other.

In the release phase, applications are deployed. Before that, system integration and tests are performed and eventual bugs are fixed by Scrum teams through scrum process.

5.2 ITV Navigation Reference Architecture

The reference architecture is a core architecture that captures the high level design for the applications of the SPL. It includes the variation points and variants documented in the variability model realized by components [15]. The iDTV navigation system variability model contains 3 variation points: language, market segment and standard. The *language* variation point has variants representing languages for the iDTV navigation system, depending on the countries which it will be used. At least one of those languages is required for each application.

The *market segment* (MS) variation point has variants representing *low-end*, *mid-end* and *high-end* market segments. Each application requires only one market segment variant. The *low-end* variant contains *this channel information* and *block programs* components. The *mid-end* variant contains the *low-end* variant and *reminder, search programs and full guide* components. The *high-end* variant contains the *mid-end* variant and *recommendation* and *personal video recorder* components.

The *standard* variation point has variants representing the ISDB [18], DVB [19] and ATSC [20] digital TV system information standards, which are realized by parser. Each application requires only one standard. The system information standards provides information about TV programs and events, like start time, duration, synopsis, etc.

5.3 ITV Navigation System Product Backlog

The iDTV navigation system product backlog contains all packages, components and interfaces from reference architecture as product backlog items. In this product backlog, the column *story detail* provides detailed requirements; *prioritization* defines the product backlog item priority; *estimate* is the product backlog item estimation size; and *sprint* is the planned sprint that product backlog item will be realized. Those columns are part of any common product backlog.

In application design sub-process, beside other activities, variants are selected to be part of specific applications. To identify the product backlog items that are part of a variant, we added the columns *variant* and *variation point*, and to identify the applications that the product backlog items will be part, we added the *application* columns representing each application of this software product line.

The product backlog is ready to be estimated and detailed by scrum teams, and to manage the iDTV navigation system SPL development using the ScrumPL process.

6. CONCLUISONS

The television sets are present in more than 97% of homes in Brazil, which represents approximately 70 million of families. The Brazilian Digital Television System is working hard to reach all these people as soon as possible. Indeed, Brazil's main objective is somehow different than any other country that has adopted Digital TV: besides improving the reception quality and increasing the amount of programs, among other benefits, Brazil wants to carry on the digital inclusion of the population.

In fact, the growth in Interactive Digital Television technology in Brazil provided an extraordinary opening to significantly broaden the role of television. Applications like T-Health, T-Government, T-Commerce, and T-Learning are some examples of value-added services offered over digital television infrastructures that start to be present in the daily life of the Brazilian population. It also provides scenarios for integrating Digital TV with other areas of knowledge such as smart environments and ubiquitous systems, allowing access to services until now not easily achievable.

The University of Amazonas, in particular the CETELI – Electronics and Information Technology Research and Development Centre – is contributing to the establishment of this new technology. We have been teaching young students and scientist, as well as proposing innovative solutions for the local market.

7. ACKNOWLEDGMENTS

We would like to thank very much the organizations and people involved in the development of this laboratory since its beginning. In particular we would like to mention Samsung Electronics and its research institute named Sidia located in Manaus. Governmental organizations like SUFRAMA, CNPq, CAPES and FAPEAM gave us support paying the costs of the laboratory and paid scholarships for our students. Most recently, CT-PIM asked our staff to prepare hardware and software courses focused on Digital TV for undergraduate students giving us the opportunity to get closer to motivated people. We also would like to thank the students directly involved in these projects, Juliano Costa, Nairon Viana, Orlewilson Maia, Lady Diana, Vandermi Silva, Antonio Santos, Ricardo Erikson, and Luciano Pinto. We are really grateful to all these people.

8. REFERENCES

[1] Rovere, R. L., " Computing in the Brazilian Amazon ", Commuications of ACM,, Vol 35, No4., 1993, pp 21-24.

[2] Despres, L. A., " Social Life and Work in Brazil's Free Trade Zone ", State University of New York Press, 1991.

[3] Srivastava, H. O. 2002. Interactive TV – Technology and Markets. Artech House, Inc. Norwood, USA.

[4] Lugmayr, A.; Niiranen, S. and Kalli, S. 2004. Digital Interactive TV and Metadata – Future Broadcast Multimedia. Springer-Verlag, New York, USA.

[5] Pagani, M. 2003. Multimedia and Interactive Digital TV: Managing the Oportunities Created by Digital Convergence. IRM Press, London, UK.

[6] Morris, S. and Smith-Chaigneau, A. 2005. Interactive TV Standards – A Guide to MHP, OCAP, and JavaTV. Elsevier Oxford, UK

[7] Piesing, J. The DVB multimedia home platform (MHP) and related specifications, in Proc. of The IEEE, jan 2006, pp. 237–247.

[8] Corcoran, P., Desbonnet, J., Bigioi, P. and Lupu, I. Home network infrastructure for handheld/wearable appliances, Consumer Electronics, IEEE Trans. on, vol. 48, no. 3, pp. 490–495, Aug 2002.

[9] Ge, C., Li, Y., Zhi, X. and Tong, W., The intelligent STB - implementation of next generation of residential gateway in digital home, in 2nd IEEE International Conference on Pervasive Computing and Applications, 2007, pp. 256–261.

[10] Soares, L. F. G., Rodrigues, R. F. and Moreno, M. F.. Ginga-NCL: The Declarative Environment of the Brazilian Digital TV System. Journal of the Brazilian Computer Society, 13(1):37-46, March 2007.

[11] Baima, R. L. Realization and Optimization of an Agent-Based Test Resources Manager. Master Thesis of the Universität Stuttgart , 2009.

[12] FIPA. Foundation for Intelligent Physical Agents. http://www.fipa.org/repository/standardspecs.html. Access in 2009 October.

[13] Isobe, T., Fujiwara, M., Kaneta, H., Uratani, N., Morita, T. 2003. Development and features of a TV navigation system. In IEEE Transactions on Consumer Electronics.Vol. 49, Issue 4, Nov, 2003. 1035–1042. DOI = 10.1109/TCE.2003.1261192.

[14] Peng, C., Lugmayr, A., Vuorimaa, P. 2002. A Digital Television Navigator. In Multimedia Tools and Applications. Volume 17, Number 1. May, 2002. 121-141. DOI= 10.1023/A:1014687823960

[15] Pohl, K., Böcke, G., Linder, F. 2005. Software Product Line Engineering – Foundations, Principles and Techniques. Springer-Verlag Berlin hedelberg. Germany.

[16] Larman C. 2004. Agile and Iterative Development–A Manager's Guide. Addison Wesley. United States

[17] Schwaber, K., Beedle, M. 2002. Agile Software Development with Scrum. Prentice Hall. United States..

[18] ABNT NBR15603-2. 2007. "Televisão digital terrestre – Multiplexação e serviços de informação (SI). Parte 2: Estrutura de dados e definições da informação básica de SI". The Brazilian Specification for System Information.

[19] European Telecommunication Standard Institute (ETSI) 300 468. 1997. "Digital Video Broadcasting (DVB); Specification for Service Information (SI)".

[20] Advanced Television Systems Committee (ATSC) A/65. 1997. "Program and System Information Protocol for Terrestrial Broadcast and Cable (PSIP)"

China's IPTV Industry, Status quo and Trend

Zhen Yu GU
Shanghai Jiao Tong University
800, Dong Chuan Road,
Shanghai City, China
86-21-34204960
zygu@sjtu.edu.cn

Lin GUI
Shanghai Jiao Tong University
800, Dong Chuan Road,
Shanghai City, China
86-21-34204960
Guilin@ sjtu.edu.cn

ABSTRACT
This paper presents an overview of China's IPTV industry. Especially the exploration of some business models based on China's reality is introduced and discussed. The bottlenecks of China's IPTV industry are analyzed. The future direction of development and then some measures are proposed at the end.

Keywords
IPTV, Business Model, Supervision Policy

1. MARKET OF CHINA'S IPTV
China has opened IPTV business since 2004, and achieved rapid development even under the discouraging and restrictive supervision policy. IPTV subscribers increased from 267,000 in 2005 to 4.7 million in 2009, subscribers increased more than tenfold [1].

IPTV combines some aspects of traditional television, radio, movie, internet and other telecommunications into one integrated form. Based on interactive broadband network, IPTV could provide internet browsing, multimedia, communications, live TV, and other services. Taking advantages of those, according to Point Topic, Inc., in 2009, the number of global IPTV users grew 47%, an increase of 10.8 million users, and the total number of users had reached 33 million. China's IPTV industry has made considerable progress following global development. Since 2008, though adverse effects caused by various unsettled key issues in the industry restructuring of telecommunications, IPTV subscribers in China has still maintained a sustained rapid growth. At the end of 2009, China mainland IPTV subscribers reached 4.7 million [1].

IPTV is operated mainly by China Telecom, China Unicom (formerly China Netcom) and China Digital TV Co. at present, which would be called "Big Three" in IPTV. In 2009, China Telecom's IPTV subscribers reached 3.62 million, in which Shanghai topped 1.01 million, followed by Jiangsu 780,000, Guangdong 710,000. The China Digital TV Co. which located in Hangzhou had developed 560,000 users. Most of them were in Zhejiang province, about 420,000 [2].

It is predicted that by the end of 2010, China's IPTV subscribers will be expected to reach 9 million [3], but compared with China's population, or even with the amount of cable TV users which is more than 100 million, there is still much room for growth. Though the market competition is fierce, China is still a big cake waiting for share for IPTV industry.

2. BUSINESS MODELS OF CHINA'S IPTV
2.1 Industry Chain
IPTV's industry chain constitutes of six participants: content providers, content and value-added application integrators, content and value-added application service providers, network and transmission operators, equipment suppliers and users. The relationship between telecommunication networks operators and content providers is the most critical [4]. Therefore, how to create cooperative business win-win models which are profitable to all participants became a real urgent need.

2.2 Business Models
IPTV is a set of Internet, telecommunications and entertainment in one of the new business, which involves the terminal PC and TV terminals. Therefore, the traditional single telecom management mode, to some extent, does not meet the needs of IPTV business development. Since 2003, after more than 6 years of operation, China's IPTV operators still face business model and ecosystem issues of the industry, such as content, business model, pricing model and the ecological chain. They are still facing the problem of sustainable development, but the operation operators are actively exploring its own development model. So far, it is hard to say some of them have found a completely suitable business model and profit model for the development of IPTV in China [5].

Similar to IPTV, a global point of view, including digital television services, the business models of all the integrative services are still in the exploratory stage. In addition to cost factors, the key of earnings must be the expansion of market scale and enhance its user ARPU [5]. The development of digital television is the case, only relying on traditional advertising profits has many problems.

In terms of companies' backgrounds, there are four kinds of IPTV business models in China:

(1) Co-operation Mode

IPTV in Shanghai, for example, is co-operated by Shanghai Media and China Telecom. The two sides carried out their services in accordance with their areas of expertise to provide. In the co-operating mode, the broadcasting side was responsible to provide content and content integration, while telecommunications were mainly responsible for the infrastructure and network of common business operations and benefit-sharing.

(2) Broadcast Group Self Operating Mode

Zhejiang, IPTV service providers have radio or television background. The IPTV business development includes the content producing, content integration, network transmission and a series of links. Self operating mode in radio and television companies can conduct live, on-demand TV program, in addition to the basic business.

(3) Telecommunication Self Operating Mode

Jiangsu and Henan, for example, the telecommunication company provides IPTV services on its own. Telecommunication network operators use their own network resources and also try to reach the front end of the profit chain. They established their own content integration divisions, to purchase content resources, then refining the content and services, to provide a full package of IPTV services.

(4) New Joint Company Operating Mode

This particular model was used in the early stage of development of Hangzhou IPTV. Initially, the joint was made in between Hangzhou Netcom IPTV and the Hangzhou Digital TV Limited, to promote the implementation of market operations. In this mode, the new entity was resulted from the market competition. The background of two partners was telecommunication and broadcasting, both had a large user of resources and network resources, rich experience in market operation.

As the IPTV industry in China is still at an early stage, so all the business models remain needing further market test.

3. BOTTLENECKS OF IPTV IN CHINA AND ITS POLICY REASONS

China's IPTV industry has, despite considerable progress, still many problems:

(1) Monotonous content

IPTV competitive edge comes from its personalized contents and interactive forms of programs. Content is one of the fundamental driving forces for IPTV and other digital new media to attract users. With the explosive growth of digital new media and increased per capita consumption of media, the day that content is king is coming. However, the main operation bodies of IPTV in China - China Telecom, China Mobile and China Unicom and other state-owned telecommunications companies do not have the competitive advantage of content. China's Internet, the copyright status of chaos and free service business form, exerts a greater pressure on the value of contents of the IPTV. IPTV operators are encountering some intellectual property issues such as excessive costs of copyrights. Current IPTV contents are mainly based on traditional television programs. IPTV content now has three main sources: radio and television, Internet and content providers. Unfortunately, without IPTV, people still can access most of these resources by other means. A survey has revealed that is the reason why most people have no motivation to pay extra to launch IPTV [6]. Innovative, attractive digital new media content is still quite lacking. In fact, our traditional TV programs can not satisfy the audience's thirst for media content, only 15% of home-brewed television programs are broadcast content [5].

For telecom companies, to provide unique content is not so easy. They are lack of operational experience in the media industry. And media production in China subject to strict supervision of radio and television administration.

(2) Supervision policy

China's IPTV market is mainly supervised by three administrative departments: the Ministry of Industry, SARFT (state administration of radio film and television) and the Ministry of Culture. Three departments have many conflictions in their supervisions since they represent different sectors of the industry with their own interests.

IPTV has the characteristics of regional market. From the competition point of view, every IPTV operator faces competition first, the local radio and television system which is SARFT's play ground.

The TV and media industry is not a free market in China. Domestic media producers in radio and television are largely supervised by SARFT. Any policy change would bring about a great impact on the development of IPTV. More seriously, in 2010, restriction on IPTV seems even harsher than before. There are always headlines of that SARFT just stopped a local IPTV business without permit. And for some reasons, the permit is very hard to get.

4. THE DEVELOPMENT TREND OF CHINA'S IPTV

Interactive features of the Internet allow easier delivery of IPTV video on demand, such as interactive games, interactive value-added services. IPTV represents the direction of triple play. The overall size of IPTV user group in China is still small. However, the market potential is immeasurable. Currently, China has more than 340 million domestic broadband users. Therefore, the large broadband user base is a solid foundation for sustainable development IPTV. How to develop the IPTV users in the broadband subscribers is the main problem that telecommunications companies and other IPTV operators in China need to urgently address.

It is a trend that China's radio and television, telecommunications and the Internet, the network infrastructure industry, are accelerating their paces from the separation to integration of the industry. In this context, after a new round of political reform and related policy adjustments in 2008, the triple play of services, including IPTV, is entering an age of rapid development.

With the increasing popularity and the expansion of the market scale, IPTV market will attract more content providers, content aggregators and value-added service providers to enter. They will provide innovative contents, and bring about a broader exploration of business models. With large-scale commercial TD-SCDMA and the completion of a new round of telecom restructuring, 3G has already entered the mass market. From the user point of view, 3G terminals can be mobile IPTV user terminals. With personalized 3G terminal, IPTV user-oriented development goals will be finally realized.

Five licensed IPTV operators have, taken the chance of 2008 Beijing Olympic Games, started to kick off IPTV movement. For the three telecom operators, IPTV will be very important to its business portfolio. While two radio and TV operators are also

transform their business models to digital new media by updating the network with two-way and interactive digital TV technology.

The barriers from IPTV policy will be gradually eased. In April 2009, the State Department released "Electronic information industry restructuring and revitalization plan". The document firmly declares support to IPTV, mobile TV and other emerging service industry.

Win-win is the only way for the development of the IPTV industry and digital TV. IPTV's media properties require operators to be market-based, network -oriented, user-centered, actively cooperated with the media, entertainment and information content services. IPTV is one of the most important businesses in the next generation network (NGN). It is also a very important business form in the future digital home. With the development of ICT, it is an inevitable trend that telecommunications networks, Internet, cable TV become converge. Development of triple play requires TV terminal and PC terminal can both connect to the Internet and cable television networks, access to the Internet and receive digital TV broadcasts at the same time. The coexistence of multiple access methods ensure an optimal manner of unicast, multicast, broadcast and two-way interactive service, to meet the needs of digital new media, to realize the integration of telecommunications and broadcasting industry value chain. IPTV and digital TV operators should abandon their prejudices and learn from each other, the strategies of development, and the operational experiences, to jointly promote the triple play, to form a win-win situation.

5. CONCLUSIONS AND PROPOSED MEASURES TO PROMOTE CHINA'S IPTV

The IPTV operators should provide more unique content and applications, consistently improve the quality of user experiences with more user-friendly interaction, to enlarge their user groups.

In addition to full use of market mechanism, Chinese government should reinforce the policies of triple play, foster competing and clustering among the telecommunications and broadcasting, digital television and other media.

5.1 Proposed Measures

1. To amend the standards of issuing "information dissemination networks audio-visual programs permit". IPTV broadcasting administration of government should increase the scope of license applicants and license number. To issue more national IPTV operator licenses. To encourage state-owned strategic investors to invest in IPTV digital content related businesses.

2. To coordinate the relevant policies of triple play, eliminate the institutional and political obstacles for triple play. Triple play will be the goal to promote the IPTV industry operations.

3. To foster competing effect, according to regulatory of IPTV market. To restrict unfair competition and monopoly (particularly administrative monopoly) behavior. To enhance technology, business, marketing and strategic level cooperation.

4. To promote the renovation of the digital TV access network and backbone network, speed up the pace of digital TV industry, and IPTV services such as Internet audio and video programs, co-enlarge digital new media business market.

REFERENCES

[1] D.Z. Bao. 2010. IPTV market is expectable. China New Telecom (Feb 2010), 5-10(in Chinese).

[2] http://www.ccidreport.com/market/article/content/408/201003/221648.html.

[3] http://iptv.lmtw.com/yj/201003/54968.html.

[4] J. Chen, JX Dong. Discussion of IPTV industry. Commercial Technology, 22-23(in Chinese).

[5] G.B. Xu.2007. Analysis of China's IPTV development mode. Telecom world (Mar 2007),55-56(in Chinese).

[6] http://media.icxo.com/htmlnews/2005/09/09/664579.htm.

Interactive Digital TV Services in India

Rajendra Akerkar
Western Norway Research Institute
Sogndal, Norway
Telephone number: +47 91685607
E-mail: rak@vestforsk.no

ABSTRACT

Interactive digital television is a potential revolution in home entertainment, enabling the convergence of many types of media. The scope of application is as grand as one's imagination. Interactive digital TV (iTV) is a system where a digital television service is delivered using Internet Protocol (IP) over a network infrastructure, which includes delivery by a broadband connection. In this paper, we will discuss the status of iTV services in India. iTV services are still in the nascent stage in the country but are being seen as a potential threat to the existing cable and satellite services.

Keywords
Interactive digital TV, internet protocol, broadband, public sector telecom, private sector telecom

1. INTRODUCTION

Television in India has been in existence for almost four decades. For the first 18 years, it spread slowly and transmission was mainly in black & white. Television has come to the forefront only in the past 25 years. There were primarily two turning points: the first in the eighties when colour TV was introduced by nationalized broadcaster Doordarshan (DD). It then proceeded to install transmitters nationwide rapidly for terrestrial broadcasting. In this period no private enterprise was allowed to set up TV stations or to transmit TV signals. The second ignite came in the early nineties with the broadcast of satellite TV by foreign programmers like CNN followed by Star TV and a little later by domestic channels such as Zee TV into Indian homes. From the large metros satellite TV delivered via cable moved into smaller towns. In recent days, Internet Protocol Television (iTV) services are gradually becoming popular in India. Many well-know telecom operators have already launched iTV services in various parts of the country.

iTV refers to TV delivered through broadband, a technology that was earlier used to provide only internet access [1, 3]. The iTV services are still in the initial stage in the country but are being seen as a likely threat to the existing cable and satellite services.

Next section presents current scene in Indian market together with reasons for slow growth and advantages of iTV. In section 3, we will discuss the key service providers in iTV market.

2. CURRENT STATUS

The companies offering iTV services in Indian include public sector telecom giants MTNL, BSNL and HFCL, and private sector companies like Airtel. Presently offered in a few cities, the iTV services are delivered in collaboration with content delivery company like Aksh Optifibres and Goldstone.

In recent times, these content delivery companies provided iTV connection without a broadband connection to encourage the dissemination of iTV among the masses. It is for sure that several large telecom players will enter into the business exploring iTV as a new revenue opportunity from their existing markets.

In the beginning, the medium of entertainment was restricted to the usual Cable operators with limited content and low pricing. But the changing technological scenario and the digitization of broadcast industry has brought a wave of change in the entertainment sector. Customers are demanding more value added-content, with digital quality at cost- effective pricing. iTV faced a slow start as people were not aware about this new experience of watching television. Over a few months, it has spread to the other cities and is now gaining momentum. iTV is also growing in popularity in the urban and rural areas due to the content and pricing. iTV is now attracting more private entrants in the space.

At the moment, the market for iTV in India is at a emerging stage, with only a handful of operators having their presence. Also in India, we have lesser percentage of Internet users compared to the other developed countries across the globe. At present, India has about 10,000 subscribers for iTV, but India is expected to be the fifth-largest iTV market by the year 2011 in the APAC region.

According to Frost and Sullivan [2], the Asia-Pacific region is likely to have over 27 million iTV subscribers by 2013, with China and India being the high-growth markets.

Any new technology takes time to reach its "point of inflection", after which it starts growing at a much faster rate. Even broadband adoption in India was very slow initially, but since 2007, it has really picked up steam. In case of iTV, with the government now giving go-ahead to broadcasters to share their channels with iTV providers, the deployment of Internet Protocol Television is all set to revolutionize and enhance television viewing experience in India. The slow but steady growth is already visible – we have observed more announcements on iTV launches in the last few months than in the last few years.

2.1 Reasons for slow growth

There are many reasons for slow growth of iTV services in India. To name a few, there is a lack of customer awareness about the benefits of the iTV service, low broadband penetration, slow Internet access speeds and costly set-top box. Most important reason why iTV never took off earlier in India was slow Internet connectivity. However, with faster broadband speed now there is every possibility that iTV will widespread soon. iTV uses a two-way broadcast signal sent through the service provider's backbone network and servers, allowing viewers to select content on demand, time shift and to take advantage of other interactive TV options. To take advantage of this, viewers will need a broadband connection and a set-top-box to send and receive requests. Traditional cable TV has the capacity to deliver hundreds of channels simultaneously to each subscriber. This creates limitations on the number of channels offered and can contribute to bandwidth shortages and quality degradation. iTV, by contrast, sends only one programme at a time.

2.2 Advantages of iTV

iTV delivery offers some advantages over cable and satellite TV providers. Unlike cable TV, iTV allows people to do a number of things. They can watch movies, play games, make telephone calls, and work on a computer and many more things, which are not possible on other platforms. In addition to these, one can do online shopping, e-learning and e-banking soon. These things have already been accepted and used by Internet and mobile users, and once iTV will capture the market, it will be widely accepted by everybody.

iTV will bring services not yet delivered such as on-demand video content, network-based DVRs, where the content is potentially stored on the network and streamed to the device wherever it might be. The so-called "long tail content" will be straight away available to consumers.

In addition, the content which is usually available in streaming media format on the Internet will be available for viewing on the television. iTV will also bring the integration of video streaming with conferencing capability, and interactive TV applications which will provide users with a much richer experience than they get today.

From the Indian consumer's perspective, iTV will allow the user to experience digital quality television with the added advantage of being a "pay per view" – service. Consumers can expect services such as Video on Demand, Video Conferencing and Interactive TV – pause, fast forward and rewind live TV or recorded content stored on the service provider's remote servers. Therefore, all parameters of comfort, cost and quality will be taken care of.

3. MAJOR PLAYERS

With the broadband penetration gaining momentum, iTV is set for a boom in India and is expected to be available in around 30 cities in the next two years. The subscribers of ITV range from residential users to government bodies to entertainment industry. ITV is currently commercially available in Pune, Mumbai, Bangalore and Kolkata and is now showing signs of gaining traction in the Indian market. Currently, iControl iTV is watched by MTNL subscribers in Delhi and Mumbai. It is also enjoyed by BSNL subscribers in 20 cities of North India, i.e. in Rajasthan, Punjab, Haryana, Jammu, etc.

Interestingly, companies like eInfochips has services and solutions offering for all stages of Video broadcast chain, i.e. Creation, Transmission, Distribution, Testing, and Consumption. eInfochips is working with many global players in the value chain of iTV which includes semiconductor companies, equipment manufacturer both for broadcasting and customer node. eInfochips has design expertise of STB (set-top box), DVR/PVR, transcoding, automated solution for video quality verifications, AV container libraries, real-time content streaming stack, etc.

eInfochips recently announced availability of H.264 HD AVC and H.264 SVC and also launched High-definition (HD) reference board design based on Texas Instruments' DaVinci media processor last year. This will reduce TTM (time to market) for system/product manufactures and designers for HD-based solution.

In Indian market, Aksh Optifibre is the pioneer in iTV technology and enjoys highest subscriber base in this industry with MTNL and BSNL. These companies enjoy the last-mile connectivity factor over the other players entering this space and hence there is no question of other players entering this segment and posing a threat. There is enough for everyone in this segment.

Airtel is the new entrant in this segment, while Reliance and some other companies have delayed their launch for the time being. Besides being a leader in iTV in China, UTStarcom is also India's leading provider of iTV and broadband infrastructure in terms of scope of deployments and total commercially-deployed subscriber lines. For the last three years, they have taken various steps to promote growth and acceptance of iTV in the country. UTStarcom enables service providers to further monetize their existing broadband infrastructure by offering iTV and complementary value-added services that can be bundled with existing voice and data services to improve ARPU and reduce overall churn. Their end-to-end iTV system can be integrated into existing broadband networks to serve as the delivery platform for revenue generating applications such as IP video surveillance, video conferencing, distance learning, digital signage, interactive voting, hospitality applications, and advanced advertising.

Aksh iControl, India's first Internet Protocol TV (iTV) service, built on the end-to-end RollingStream iTV solution framework provided by UTStarcom, and being offered through BSNL and MTNL, allows end-users to pause, rewind and fast-forward, as well as record, live TV. However, in order to differentiate itself from other video services available in India – or even in other global markets – this interactive iTV offering will enable subscribers to message and talk with one another live through video communications tools on their television, as well as interact live with their favourite shows through polling and karaoke. iControl also offers a time shift feature that allows viewers the ability to view missed programming that was aired in the past week. The Video-on-Demand feature brings a library of more than 200 Hollywood, Bollywood and regional content by pressing only a few buttons on the TV remote. With another feature, CAS allows users to subscribe and de-subscribe a certain TV Channel or Channel group using just your TV remote and TV without any need of calling a call centre or suffering a long iTV delay or filling a form for the selection of channels. A-Tube is India's

primary innovative video yellow pages, which provides a variety of information ranging from lifestyle to video resume. A-Shop is the viewers' outlet to buy and sell various products in a virtual marketplace. Similarly, Bharti Airtel is also providing a lot of value-added and innovative services on its iTV platform.

4. SUMMARY

The response from iTV users has been positive in Indian market. However, to make iTV more popular, the service providers should lay stress on the quality of services and strengthen the customer care department, which is in a very poor state at present. Moreover, service providers need to look beyond instant revenue opportunities to understand the durable worth of iTV as a carrier distribution platform, over which several consumer communication and entertainment services can be offered concurrently.

5. REFERENCES

[1] Srivastava, H. O. 2002. Interactive TV – Technology and Markets. Artech House, Inc. Norwood, USA.

[2] Frost and Sullivan 2009. Asia-Pacific IPTV Updates http://www.frost.com/prod/servlet/frost-home.pag

[3] Lugmayr, A.; Niiranen, S. and Kalli, S. 2004. Digital Interactive TV and Metadata – Future Broadcast Multimedia. Springer-Verlag, New York, USA.

[4] Pagani, M. 2003. Multimedia and Interactive Digital TV: Managing the Opportunities Created by Digital Convergence. IRM Press, London, UK Series. ACM Press, New York, NY, 19-33. DOI= http://doi.acm.org/10.1145/90417.90738

Ginga Edutainer: A Framework for the Development of T-Learning Applications on Ginga – The Brazilian Digital Television

Newton Nyamasege Marube, Daniel Campelo Tavares
Federal University of Rio Grande - FURG
Center of Computational Sciences - C3
Campus Carreiros, Av. Italia km 08, 96201-900
+55 53 3233 6582
{nyamasege, prof.danieltavares}@gmail.com

ABSTRACT
The aim of this study is to propose a framework for developing educational applications for Digital Television, thus simplifying the process of developing applications by abstracting the execution platform. The arrival of digital TV in Brazil offers a range of possibilities in the field of education. The availability of television sets in 95.1% of households in Brazil indicates that studies on television as a form of teaching and learning can be adapted to this medium. The proposed tool incorporates entertainment and education and aims to make the task of producing educational software simpler, and consequently reducing the time needed to develop applications that can be used to educate and promote digital inclusion in the country. When implemented, the project will integrate interactivity in education, creating educational applications that can motivate the user in an environment of teaching and learning.

Categories and Subject Descriptors
D.3.3 [**Programming Languages**]: Language Constructs and Features – abstract data types, polymorphism, control structures. H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems—Artificial, augmented and virtual realities; J.7 [**Computers in Other Systems**]: Consumer products; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – graphical user interfaces, input devices and strategies, interaction styles, screen design, training, help and documentation, user-centered design.

General Terms
Design, Human Factors, Standardization, Languages.

Keywords
T-learning, Interactive Digital TV, Ginga, Middleware, Frameworks.

1. INTRODUCTION

The arrival of digital TV in Brazil took place in June 2006 when the Brazilian government announced the selection of ISDB-T, developed in Japan, as the basis for the development of the system adopted in Brazil ISDB-TB, also called SBTVD. In the Brazilian version, technologies developed by the Catholic University of Rio de Janeiro (PUC-Rio) and the Federal University of Paraíba (UFPB) were added. The start of SBTVD transmissions occurred on 02 December 2007, in Sao Paulo. Peru, Argentina, Chile and Venezuela have announced the adoption of the Brazilian standard of digital TV ISDB-T. [1]

The television is a largely available device that most individuals are very accustomed to. The National Household Sample Survey – carried out by IBGE (Brazilian Institute of Geography and Statistics) in 2008, found that 95.1% of Brazilian households possessed at least one television.

Table 1. The number of Brazilian households and Brazilian Household with TV (%) from 2004 to 2008

%	2004	2005	2006	2007	2008
Television	90.3	91.4	93	94.5	95.1
Households*	51753	53114	54610	55770	57557

***households in thousands Source: PNAD (IBGE)**

People usually trust what they receive through the television set and the quality of the content assures a richer experience for the user. Television can perfectly cope with generating very rich content and high levels of motivation for the user, these are among the factors that work in advantage of the television as an educational tool. [5]. The main advantage of digital TV is that hundreds or thousands of programs can be offered in parallel, which means that personalized applications can be offered to specifically defined user groups. With respect to education, this means that it is possible to customize the delivered educational content according to the needs of different classes of learners such as age, learning styles, level of formal education or personal interests.

A few years ago when Digital TV was conceived, the range of possibilities was limited by technology, mainly the unavailability of a return channel that is essential for feedback and disk space limitation in set-top-boxes for optional storage of the digital content for later re-viewing.

2. DIGITAL INTERACTIVE TV IN LEARNING

2.1 What is Digital TV?

Digital Television (DTV) is an advanced broadcasting technology that has transformed television viewing experience. DTV has enable broadcasters to offer television with better picture and sound quality. It also offers multiple programming choices, called multicasting and interactive capabilities. [15]

Digital transmission in Brazil is gradually spreading from city to city. The analogue transmission will continue to occur, simultaneously with the digital, for a period of 10 years until 29/06/2016. As from Jul/2013 only digitally transmitted channels will be granted transmission licenses. Today, a large number of TV channels are transmitted digitally via satellite, cable and terrestrial (over - the-air). Most TVs can only receive analog signal therefore require a set-top box to receive the digital signal using a digital tuner and then convert the signal to an analog format for display on the TV sets. The set-top box is connected between the TV and the signal source. Although some television makers already have TV models marketed with an integrated structure of the STB.

.

2.2 What is interactive TV?

Interactive Digital Television is a two-way interactive service provided through television, enabling audiences greater choice, control, and customization over their viewing experience. Through new digital technologies, interactive television services such as video-on-demand, enhanced television, interactive program guides, and email are just some examples of the next generation of digital programming over cable, satellite and terrestrial broadcast television. [6]

2.3 Interactivity through digital TV

2.3.1 Navigation

Because of the large number of channels available through interactive TV, an electronic program guide (EPG) is essential for the viewer to find and select programs. The service provider is responsible for this guide, and has the liberty to organize it to suit his own needs. Selecting a program through an on-screen EPG is the most commonly used form of interactivity by form of remote control.

2.3.2 *Enhanced* TV

A number of options are available for enhancing the viewing experience through the use of interactivity of existing TV programs. Selecting options from a menu to get to additional information can be achieved by pressing one of the four colored buttons on the remote control. In 2009 during one soap opera, a local broadcaster offered interactive services that detailed previous episodes and cast information. The same services could also be seen during the Brazilian Carnival.

2.3.3 Channel independent interactive services

Interactive services are also available independent of TV video channels. These services tend to be accessible via the interactive services button on the remote control or an infrared keyboard when the interactive services icon is displayed on the screen. The child is taken to the interactive area with various activities to interact based on a popular children's program like "Sitio do Pica Pau Amarelo" or "TV Xuxa." The arrow keys or the four color coded buttons on the remote control are used to select items and navigate around the screen. The method used is really an advanced form of the method used to access teletext. Upon selecting a particular option, a module is downloaded into the very limited memory of the set-top box. This usually takes a few seconds, as the viewer has to wait until that particular option becomes available from the broadcaster.[4]

2.3.4 Interacting through the return channel

Another way of interacting is through the return channel. In the case of satellite and terrestrial transmission this is through a telephone connection to the set-top box. For digital cable this is through the cable network.

Utilizing the return channel enables viewers to respond to questions through yes or no or multiple choice questions and also enables viewers to vote. This same method can provide access to TV Internet – a more limited version of the Internet that has been modified for easier viewing on the TV. It is also possible to send emails usually without the ability to attach documents. An infrared keyboard can be used to type in messages instead of the remote control. Data from the last National census shows that less than 30% of Brazilian homes have internet access. The government has plans to provide internet coverage countrywide although many consider this a long shot.

2.4 What is t-learning?

The term "t-learning" was adopted as an abbreviation to mean TV-based interactive learning. T-learning is about having access to interactive video-rich learning materials especially at home through a TV.

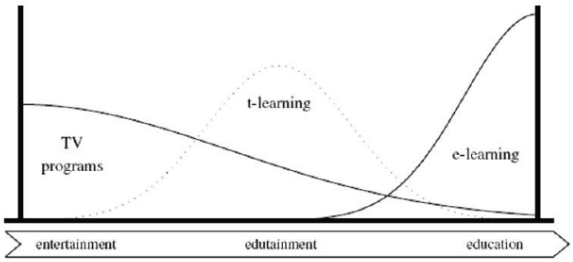


Figure 1. The scope of t-learning: between pure entertainment and formal education. [12]

While e-learning currently tends to be used to mean learning over the Internet using a personal computer, it could be used to refer to any form of learning using a connected digital electronic device. Thus, t-learning is really a subset of e-learning - with access through a home-based TV. T-learning could significantly improve the learning experience in a way that Internet-based e-learning is currently unable to. [4]

2.5 Role of interactive TV in learning and digital inclusion in Brazil

Developments in interactivity are raising awareness that the TV can start to be used actively rather than just a passive mode. Broadcasters and service providers are still experimenting with the interactive services. Despite, slow developments towards personalized TV, the time is right to start focusing more attention on this area as a means of creating new, more personalized learning opportunities at home. This is likely to be a more fruitful way of widening access and participation in learning in the medium term scenario (within the next five years) rather than focusing on just the interactive TV available through broadcast tv. Value can be aggregated to the learning experience when these programs are available in a video-on-demand mode rather than in a broadcast mode – assuming that all the enhanced features are available.

Unless there are some significant advances in the technology, broadcast or scheduled TV interactive learning opportunities are likely to be always limited. Since internet access for all still remains a dream, Digital TV can rise up to fill the digital divide seen in Brazil on present day. Interactive TV can reach almost every home in the country at lower costs thereby helping to have digital inclusion countrywide.

3. GINGA MIDDLEWARE

Ginga is the name of Open Middleware for the Brazilian Digital TV (SBTVD). Ginga is composed of a set of standardized Brazilian technologies and innovations that make the specification advanced and the best middleware solution suiting the requirements of the country. [1]

Ginga open middleware is divided into two main integrated subsystems that enable the development of applications following two different programming paradigms. Depending on the functionality of the design of each application, one paradigm is more appropriate than the other. These two subsystems are called Ginga-J (for Java applications) and Ginga-NCL (for declarative NCL applications). The links above have specific information about the two systems.

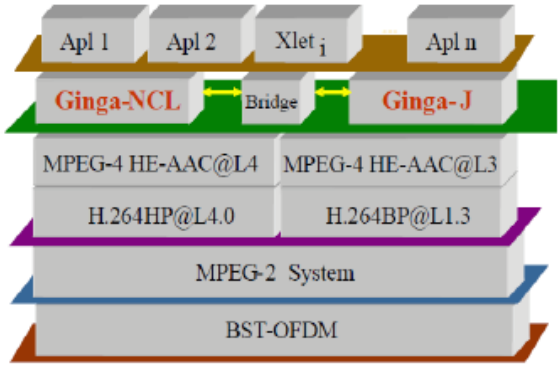


Figure 2. Ginga Architecture [8]

3.1 Ginga – NCL

Ginga-NCL is the Ginga subsystem responsible for the view of NCL documents and was developed in order to provide an infrastructure for submission declarative applications written in the NCL language. NCL is an XML application language with facilities for specifying aspects of interactivity, timing-space between media objects, adaptability, support for multiple devices and support the production of live interactive non-linear.

The Nested Context Language (NCL) is a declarative language for authoring hypermedia-based documents on the conceptual Nested Context Model (NCM) and was developed using a modular structure, following the principles adopted by the W3C. Thus, the modules for the specification of connectors and templates for composition, called XConnector and XTemplate respectively, can be incorporated into other existing languages, such as XLink, XHTML and SMIL, used for document authoring on the Web.

3.2 Ginga- J

Ginga-J (execution machine) is a logical subsystem of the Ginga system that processes procedural applications (Java Xlets). A key component of the application environment is the procedural mechanism for implementing the procedural content, which is based on a Java virtual machine.

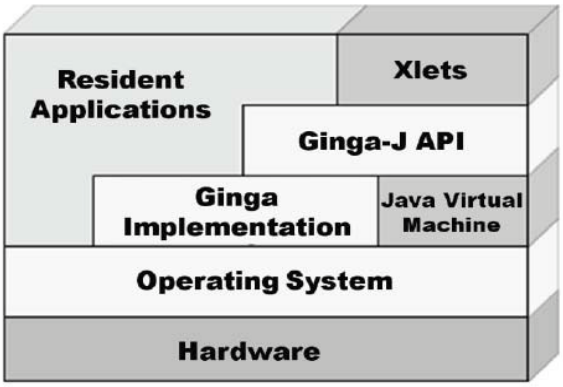


Figure 3. Ginga-J architecture and execution environment [ABNT/CEE-85 3º PROJETO 85:000.00-006/4 MAR 2010]

The resident applications can be implemented using non-standard functions provided by the operating system of the Ginga device, or by a particular implementation of Ginga. Resident applications can also incorporate the functionality provided by the standard Ginga-J API. Transmitted applications (Xlets) should always use standard API provided by the Ginga-J.

4. METHODOLOGY

To achieve the proposed objectives, the work was divided into two parts. Firstly, several studies have been carried out so as to have be a better understanding of existing technologies, of how Ginga is structured, and how T-learning applications will work and what characteristics will be passed to the new framework. In the second part, the implementation of the work will take place, where the framework will be developed and the functionalities proposed implemented.

5. GINGA EDUTAINER FRAMEWORK

Ginga Edutainer is a Digital TV application development framework proposed in this paper. The main aim is to provide a structure that facilitates the development of content that is not only educative but also entertaining for the Brazilian Digital TV.

The purpose of creating a software framework for application development is to avoid the monotony in implementing common tasks over and over again every new application is produced. Valente (2005) present a similar approach, but focuses on the reuse of software components for computer game development. Software development auxiliary tools has great importance and the creation of frameworks that allows greater code reuse and reduces the needing to rewrite code for common tasks may help to make the process quicker and more intuitive.

The Ginga Edutainer technical framework will consist of a list of features identified within the framework and for each feature there will be an agreed definition of the scope and purpose of the feature, a list of applicable standards and specifications, general guidance on creating, exhibiting and consuming this. The framework aims to support creation of multiuser t-learning applications.

6. RELATED WORK AND DISCUSSION

There is no published work on T-learning on the Ginga platform but Diego *et al.* [11] report having created a Ginga Game framework that is a framework for construction of Games. In Europe however three projects can be cited as related to this project, Kamila Olševicová *et al*. [2] present a T-learning approach enhancing video with active content. Lytras M. *et al* [5] talk about interactive TV and e-learning and try to model the convergence point of the two. While Arias *et al*. [12] present ATLAS: a framework to provide multiuser and distributed t-learning services over MHP. Francesco Bellotti *et al.* [13] talk of designing a Constructionistic Framework for T-Learning.

The Ginga Edutainer framework promises to revolutionize learning as we know it. Helped by the fact that every home in the country has a Television set education has the opportunity to reach every household. The benefits are uncountable. With the framework, language learning applications can be created and distributed countrywide through television. The framework will be able to produce applications that integrate on-demand content with ongoing broadcast.

Currently the execution of this project is delayed because the approval of JavaDTV specification is in voting stage and the binaries are therefore unavailable currently and will be out in the second semester of the year. Besides this setback other problems face a full scale implementation of T-learning, although it is probably accepted that the television in its traditional format is a very powerful medium, the body of research into its role for learning is rather more limited. Research has tended to focus on the impact that TV makes on individuals. There appears to be limited research into the importance of informal learning as a means of drawing people into formalized learning. Despite some evidence to suggest that more people wish to learn from home, there appears to be limited work into understanding the conditions and requirements that are needed in order to make the home a conducive learning environment.

Understanding the role of interactivity is a very complex process with most, but limited, research focused on interactivity in computer-based environments. Unsurprisingly, there has been little research into using interactive TV for learning purposes.

The costs of accessing learning through interactive digital TV should also be considered. The framework will only be responsible for the creation of applications, needing ways to distribute content. This requires funds as minimum as that may be.

7. CONCLUSIONS

After completing research studies which comprise of the first part of the work and acquiring the knowledge required, the second part will be initiated with the intention of implementing the Ginga Edutainer framework. The development of educational applications for Digital TV in Brazil, using Ginga middleware (Ginga-J or Ginga-NCL) is possible. The creation of Ginga Edutainer aims to make the process of creating applications for Digital TV simpler, providing an environment that abstracts the execution platform leaving the creators to focus only on the creation process and not the production.

8. ACKNOWLEDGMENTS

Our thanks to PPGMC for allowing us to user their establishments to perform this research.

9. REFERENCES

[1] Fórum do Sistema Brasileiro de TV Digital Terrestre http://www.dtv.org.br/ Accessed on 19 March 2010

[2] Kamila Olševicová, Hana Rohrová, Jaroslava Mikulecká - T-learning Approach: Enhancing Video with Active Elements, 2nd European Conference on Technology Enhanced Learning

[3] Manuel Damásio T-Learning and Interactive Television Edutainment: the Portuguese Case Study

[4] A study into TV-based interactive learning to the home - Peter J. Bates, pjb Associates, UK 5 May 2003

[5] Lytras M, Lougos C, Chozos P, Pouloudi A. Interactive television and e-learning convergence: Examining the potential of t-learning. Proceedings of the European Conference on E-Learning. Academic Conferences International: Reading, 2002

[6] Based on a definition supplied on PBS website at http://pbskids.org/cyberchase/itv.html Accessed on 19 March 2010

[7] IBGE, Diretoria de Pesquisas, Coordenação de Trabalho e Rendimento, Pesquisa Nacional por Amostra de Domicílios 2007-2008.

[8] ABNT/CEE-85 3º PROJETO 85:000.00-006/4 MAR 2010 Televisão digital terrestre — Codificação de dados e especificações de transmissão para radiodifusão digital — Parte 4: Ginga-J — Ambiente para a execução de aplicações procedurais

[9] Valente, L., 2005. GUFF: Um Framework para desenvolvimento de jogos, Master's Thesis - Universidade Federal Fluminense – Instituto de Computação. Available at:

http://guff.tigris.org/docs/Thesis05-pt.pdf [Accessed on 21 March 2010].

[10] Ginga Digital TV Middleware Specification http://www.ginga.org.br/ Accessed on 21 March 2010

[11] Diego Cordeiro Barboza, Ginga Game: A Framework for Game Development for the Interactive Digital Television, VIII Brazilian Symposium on Games and Digital Entertainment Rio de Janeiro, RJ – Brazil, October, 8th-10th 2009

[12]José J. Pazos-Arias et al, ATLAS: a framework to provide multiuser and distributed t-learning services over MHP, Software—Practice & Experience archive Volume 36 , Issue 8 Pages: 845 - 869 July 2006 ISSN:0038-0644

[13] Francesco Bellotti et al, Designing a Constructionistic Framework for T-Learning

[14] A T-learning Courses Development and Presentation Framework F. Bellotti, S. Vrochidis, E. Parissi, P. Lhoas, D. Mathevon, M. Pellegrino, G. Bo and I. Kompatsiaris IEEE MULTIDISCIPLINARY ENGINEERING EDUCATION MAGAZINE, VOL. 3, NO. 3, SEPTEMBER 2008

[15] What You Need to Know About Digital TV http://www.dtv.gov/index.html Accessed on 21 March 2010

Interactive Virtual Worlds in Brazilian Digital Television

Daniel Campelo Tavares, Newton Nyamasege Marube, Silvia Silva da Costa Botelho, Nelson Lopes Duarte Filho
Federal University of Rio Grande – FURG
Center of Computacional Sciences - Campus Carreiros, Rio Grande – RS 96201-900
+55 53 32336623
{prof.danieltavares, nyamasege, silviacb.botelho}@gmail.com, dmtnldf@furg.br

ABSTRACT

This work describes a methodology for research involving technologies in interactive virtual 3D environments and the Brazilian digital television system, during the conduct of the study we will discuss programming techniques for digital television, description of virtual environments in X3D and the Ginga middleware architecture. The objectives of this research is to propose a solution in modules that provide portability of virtual environments described in X3d for the Brazilian digital television system.

Categories and Subject Descriptors

H.5.1 [**Multimedia Information Systems**] D.4.7 [**Organization and Design**]: Interactive systems

General Terms

Design, Human Factors, Standardization, Languages, Verification.

Keywords

Digital TV, Ginga, X3D, Interactivity, Virtual Reality

1. INTRODUCTION

Television is the best selling appliance in the world and along with it comes a new technological and behavioral phase, that is, television is no longer a passive medium. The interactive digital TV brings viewers the opportunity to shop, participate in polls, send emails, check bank account balance, customize programming and much more just at the touch of the remote control. Another feature offered is T-Commerce, a name for e-commerce done via interactive TV.

Digital TV in Brazil spread significantly in mid-2006 and 2007. The first official broadcast took place in December 2007.

According to Oliveira and Albuquerque (2005), Brazil is in the process of deployment of Digital TV (DTV), which can be understood as an evolution of analog TV. Among other features, DTV will allow a significant improvement in quality video and audio broadcast, including mobile devices and transmission of various programs by the same broadcaster multiprogramming [1].

The main debate involving digital television in Brazil was mainly based on the DTV standard, which could be adopted.

In Figure 1 we have a vision of the specifications adopted by many countries:

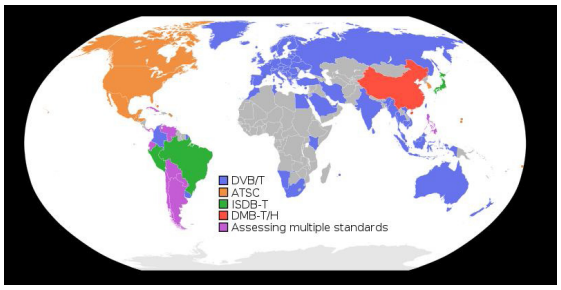


Fig 1. World Specifications [1].

The ISDB-Tb, an adaptation of the Japanese standard ISDB-T (*Integrated Services Digital Broadcasting Terrestrial*) together with technologies developed by research in Brazilian universities was defined as the official standard. The same has already been adopted in several South American countries; Peru, Chile, Argentina and Venezuela. A DTV standard is formed by sets of definitions and specifications that deal with applications, middleware, audio compression, video compression, transmission and modulation .

3D Virtual Environments, enable the insertion of the User into a virtual scenario created by software technologies, allowing the use of interactivity, navigability and immersion, characterizing the use of virtual reality.

Virtual reality is an advanced technology capable of providing the User the possibility of exploring a virtual environment in a manner similar to the real environment. In possession of the latest technology in sensors and devices, the User can navigate through the environment, finding and viewing objects at different angles, as well as acting in the scenario in which he is immersed [3].

Among the standard description of 3D environments, the format X3D is an open standard for distributing 3D content. The X3D is not an Application Programming Interface (API), nor a file format for geometry exchange. The format combines both geometry and descriptions of behaviors snapshots in a single file. The core of

the X3D specification is continually being developed by the X3D Specification Working Group [4].

It can be argued that the technology of the Brazilian DTV provides support for interactivity. The X3D standard also presents procedures for the description of interactive 3D environments. Due to the similarity in interactive features, the question arises: why not port the of navigation interactivity characteristics of X3D for the Brazilian DTV? Due to the newness of such technology in the country, there are still no studies addressing the issue of portability associated with virtual environments and the standard adopted by the Brazilian DTV.

This paper introduces the concept of virtual environments for interactive digital television. the key features associated with the two standards applicable to provide interactivity and 3D visualization will be raised.

Due to the primitive architecture of the digital image in the current decoders (Set Top Box), this work will point the directions of the requirements for the basic architecture for the support of 3D applications.

The proposed architecture will then be presented being capable of mapping the equivalence and the necessary resources in GINGA-J API and the X3D language, seeking the integration of technologies. Finally tests and analysis will be presented to validate the proposal.

2. BRAZILIAN SYSTEM FOR DIGITAL TELEVISION

In July 2006, the forum of the Brazilian system of digital terrestrial television was created by Decree 5820. Among other tasks, it is for this forum to consider the technical aspects of the generation, distribution and reception of high-definition digital television systems, (HDTV).

To meet certain technical issues, in April 2007, the ABNT (Brazilian Association of Standards and Techniques), installed a Special Study Commission on Digital Television (ABNT / EEC). This commission has developed various standards, such as:

ABNT NBR 15601 - Transmission System;

ABNT NBR 15602 - Coding of audio, and multiplexing;

ABNT NBR 15603 - Multiplexing and service information (SI);

ABNT NBR 15604 – Receivers;

ABNT NBR 15605 - safety Topics;

ABNT NBR 15606 - Data coding and transmission specification for digital broadcasting;

ABNT NBR 15,607 - Interactivity channel;

ABNT NBR 15,608 - Operation Guide;

ABNT NBR 15609 - Test Suite;

ABNT NBR 15,610 - Testing for receivers.

2.1 Ginga

GINGA, the Brazilian DTV system, has two lines of research and development (Ginga-NCL) and (Ginga-J). Before we can detail these two systems in full, we discuss the Ginga middleware, consisting of a set of hardware and software aimed at decoding systems generated in Ginga-NCL and Ginga-J to DTV.

In terms of distributed computing, middleware is a program that mediates between other software being used to hide communication protocols, platforms and peculiarities of the operating system from the programmer, allowing device independence.

In Figure 2 we describe the modules that make up the architecture of the middleware [5].

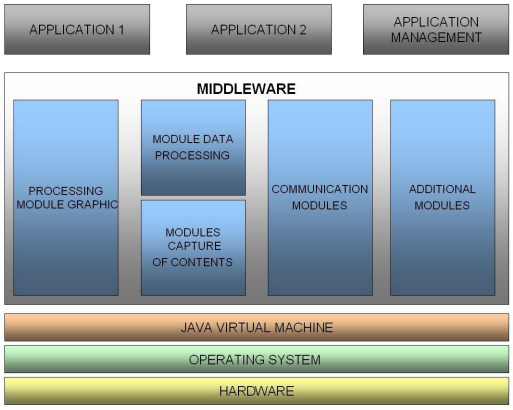


Fig 2. Structural Modules of the middleware [5].

The access terminal of interactive DTV system is represented by a layered architecture where each layer offers services to the higher layer and uses services offered by the underlying layer. Thus, the applications running on DTV make use of services of a middleware layer. [5].

Ginga is divided into two major interconnected subsystems that allow the development of applications following two programming paradigms.

Ginga-NCL was developed at PUC-Rio (Catholic University of Rio de Janeiro). Sets a presentation environment for declarative applications written in NCL (Nested Context Language) [2].

Ginga-J was developed by UFPB (Federal University of Paraíba). It provides an infrastructure for running applications based on Java.

Figure 3 characterizes the integration and structure of each language:

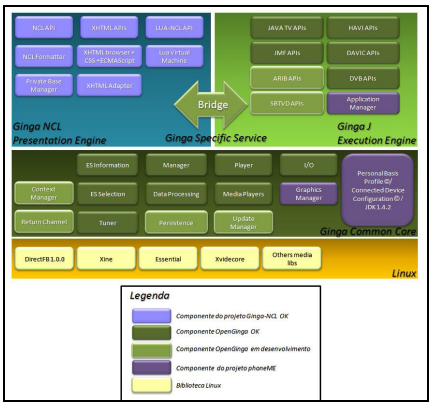


Fig 3. Integration Modules Ginga (NCL and J) [5].

In this research the programming paradigm chosen was the procedural Ginga-J based on the fact that it provided an ease of control over the code, execution and flow control.

2.2 The Ginga-J (javaDTV)

The Java API DTV adopted by ginga-j provides a lot of packages for the development of applications for digital TV, the language specification was published on July 17, 2009 as version 1.2.1, and the executables for using the API are not yet available to-date.

The structure of the source code of applications is called Xlet. The Xlet has 4 states and methods are called according to the changes of the state. When Xlet is loaded it's Paused and Active states repeats, until it is Destroyed, from where it cannot recover. Transition to the next state is not completed until the method executed according to the state change is complete. [6]

The process of the life cycle of applications is shown in Figure 4.

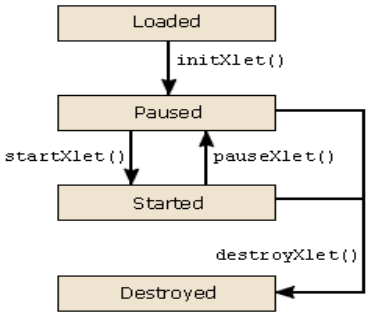


Fig 4. Xlet Life-cycle in Java-TV [6].

3. THE X3D

key interaction features with objects in a X3D scene can also be described. For this purpose it is important to understand the concepts of a X3D event.

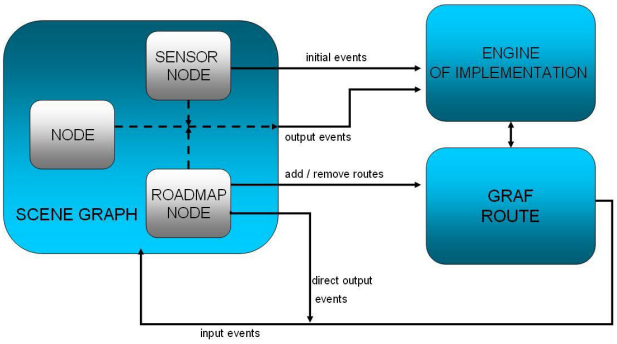


Fig 5. Events in X3D [7].

An event created from a node will spread to another node if there is a path defined between them.

Another important concept is that the fields of one node work as an attribute and at the same time as a function. In the specification of the X3D components the fields of a node are marked as inputOnly, outputOnly and inputOutput [7].

X3D offers features that allow the drawing of geometric figures, text, inserting pictures and manipulation of animated figures, as well as scenarios, objects and 3D worlds.

4. A PROPOSAL FOR VISUALIZATION AND INTERACTION OF 3D ENVIRONMENTS IN THE BRAZILIAN DIGITAL TV

Chapter IV will address a study on the technology required for implementation, and the architecture model proposed, as well as the methods needed to implement the proposal.

The methodology used in this article is an analysis of programming techniques for digital television and virtual environments. Researching the basic drawing primitives in X3D and making the relationship with packages, classes and methods needed for portability on DTV. Enabling perhaps a new feature for the Brazilian DTV.

4.1 The X3D-GINGA Architecture

The X3D-GINGA execution machine is the end product of our work and serves the purpose of the porting virtual environments described in X3D into the Brazilian DTV system. The modules that make up the execution engine are responsible for standardization of 3D content.

The model consists of three modules used during its implementation.

- The reduction module is intended to simplify the content so that the application is not too heavy due to hardware limitations imposed by the Ginga middleware;
- The rendering module prepares the content before it is transformed into an application acceptable by the DTV standard;
- The conversion module is responsible for standardization and adaptation of content for DTV, making the content readable and compatible with the application management middleware.

The flow and execution and communication can be shown in Figure 6.

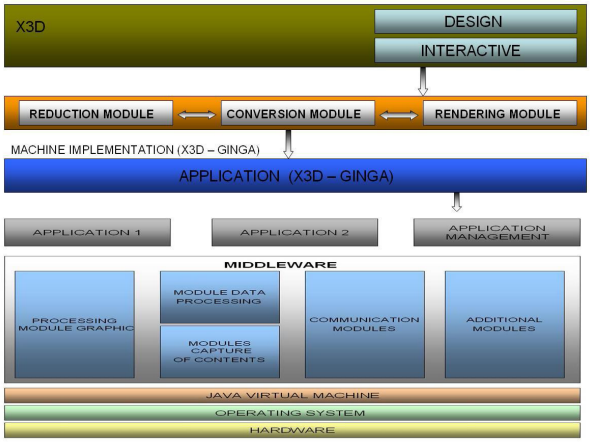


Fig 6. Conceptual model for X3D-GINGA architecture

4.2 X3D – Ginga Execution Machine

Firstly, the study of the behavior and needs for the use of X3D language needs to be carried out. Interactive virtual environments are formed by a set of primitive interconnected, related and synchronized forms.

We can say that implementation of an execution engine responsible for portability between different environments and technologies, requires a mapping of basic commands of the technologies, in order to provide an input / output library of our system.

A study in the javaDTV specification led to the use of the API packages LWUIT incorporated in the javaDTV specifications, in order to facilitate the interface handling, design of shapes and animations.

Packages that possess classes and methods needed to create the classes that make up the X3D-GINGA execution machine are:

- The package com.sun.dtv.lwuit.Graphics used for the development, design of objects and geometrical figures;
- The package com.sun.dtv.lwuit.animations, has classes for both animation of forms;
- The package com.sun.dtv.lwuit.geom contains classes related to geometry and local calculations.

Using the forms of X3D content and packages necessary for the creation of corresponding classes, we created the X3D-ginga execution machine.

5. TESTS

Application development (X3D-GINGA) started from the complete relationship of tasks from both technologies.

Figure 7 describes the mapping of common features in the languages and how it was done the mapping.

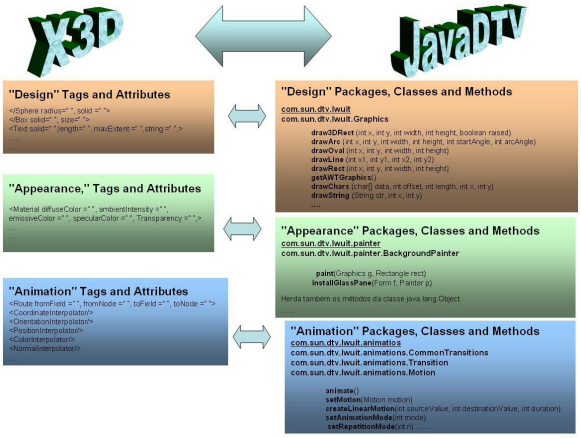


Fig 7. Relationship characteristics and procedures in languages.

Since the API specification is the only available resource, implementation will occur in the near future.

6. CONCLUSION

Hence technologies involving the Brazilian system of DTV were studied, along with the description of 3D virtual environments thereby emphasizing the technologies (Ginga middleware, ginga-j and X3D). This study brought into light an unpublished research proposing the conversion and porting of interactive 3D environment settings into digital television.

At the end of the study evidence was obtained that there are viable possibilities for the implementation of virtual environments, i.e., files with 3D extension support and portable to the Brazilian digital television system.

7. REFERENCES

[1] J. Oliveira, E. C. R. de; Albuquerque, C. V. N. (2005) "TV Digital Interativa: Padrões para uma nova era", In: V Escola Regional de Informática RJ/ES da Sociedade Brasileira de Computação, 2005.

[2] Pontifícia Universidade Católica do Rio de Janeiro LAVID, Laboratório de Aplicações de Vídeo Digital. Ferramentas. [Online]. Available at http://www.telemidia.puc-rio.br/pt/index.html.

[3] L. C. A. Rinaldi, et al. "Fundamentos e Tecnologia de Realidade Virtual e Aumentada" . SBC. 2006.

[4] Web3D, "X3D Public Specifications". [Online]. Available at www.web3d.org.

[5] Ginga Digital TV Middleware Specification. [Online]. Available at http://www.ginga.org.br/.

[6] Java ME Technology – Java DTV [Online]. Available at http://java.sun.com/javame/technology/javadtv/.

[7] D. Brutzman & Leonard Daly, X3D – Extensive 3D Graphics for Web Authors, Morgan Kaufmann, 2007.

An Application Model for Digital Television in e-Learning

Paloma Maria Santos, Marcus de Melo Braga, Marcus Vinícius Anátocles da Silva Ferreira, Fernando José Spanhol
Post-Graduate Program in Knowledge Management and Engineering – Federal University of Santa Catarina
Campus Universitário, Trindade, Florianópolis/SC, Brazil.
pmariasantos@yahoo.com.br, marcus@egc.ufsc.br, marcus.ferreira@unisul.br, spanhol@led.ufsc.br

ABSTRACT
The current E-learning applications for use in different Digital TV standards are still, in general, very incipient. Based upon studies on several applications available in the American, European and Japanese scenarios, we have identified their main features and contributions in order to develop a model that aims to take advantage of those applications for the Brazilian System of Terrestrial Digital Television (BSTD-TV). The methodology adopted includes an analysis of existing strengths, weaknesses, opportunities and threats (SWOT) in the Brazilian market as a basis for the proposed model. Finally, it presents critical success factors and some recommendations for the application of this technology in the e-Learning field.

Keywords
Interactive Digital Television (iDTV), e-Learning, Distance Education.

1. INTRODUCTION
Television, as we know it, is going through a major worldwide renewal spurred by the gradual implementation of the digital system. This change has already induced a deeply changed scenario, imposing new business models and new habits for their users. Television is nowadays the broadest means of communication. Winck [1] presents an interesting report that, there are more TV viewers today than the sum of readers in all Western History: around four billion people. For Zuffo [2], the Brazilian open television model is an instrument of national cohesion, being accessed by more than 90% of the urban population.

The Brazilian System of Terrestrial Digital Television (BSTD-TV) was created with the purpose of guaranteeing digital inclusion through interactive resources that enable Internet access and the democratization of access to information. iDTV will bring about a real revolution, since that new technology will offer resources that are as yet unknown to users. In this paper , we will evaluate potential contributions of iDTV, considering its features and peculiarities as an instrument for Distance Education initiatives by means of iDTV, also named t-Learning.

Section 2 explores the main features and concepts of iDTV and introduces some of its main applications. In Section 3, the methodology used in this study will be presented. Section 4 approaches the concept of *t-Learning* and its main peculiarities, emphasizing the currently available technologies. In Section 5, we propose applications of iDTV resources for e-Learning and, after analyzing a set of cases, a SWOT analysis is developed and presented together with the main Critical Success Factors (CSFs) and recommendations based on iDTV current scenario, as well as the advantages to be derived therein. Section 6 introduces the proposed model for iDTV application in e-Learning, its potentialities and limitations and, in Section 7, final considerations and suggestions are presented for further.

2. INTERACTIVE DIGITAL TELEVISION (iDTV)
Interactivity along with portability and mobility enabled by BSTD-TV are the competitive advantages in relation to the other existing systems: the American System (ATSC-T), the European System (DVB-T), and the Japanese System (ISDB-T).

Many authors establish possible interactivity levels from iDTV [3, 4, 5]. Lemos [4] identifies five interactivity levels that start at a minimum level – that is, reduced interaction with the TV set –, until a level where the user can effectively interact with the content. Montez and Becker [5] add three more interactivity levels to the ones proposed by Lemos, starting at the interaction of the TV viewer with a program created by the TV channel, up to a higher level, where the TV viewer contributes with content creation.

Interactivity can be made available, basically, in three ways [6]: locally (without return channel); interactivity by means of a unidirectional return channel; and interactivity by means of a bidirectional return channel. The higher interactivity levels demand bidirectional return channels and implicate higher implementation costs.

The focus of this paper is on the utilization of iDTV interactivity resources for e-Learning. For this reason, only services and applications that can be used in activities developed for this field will be included here.

3. METHODOLOGY
The methodology adopted to elaborate this study can be summarized in Figure 1. Taking as a starting point the current situation in the world scenario, 21 cases of utilization of *t-Learning* technologies were analyzed, from Europe, Japan and the United States. Then, a SWOT analysis was developed, identifying the *Strengths*, *Opportunities*, *Weaknesses* and *Threats* of the scenario of Digital TV implementation in Brazil. Based on the results of SWOT analysis, we identified the main Critical Success Factors (CSFs), and offered recommendations that should be taken into account in the BSTD-TV model, and also in the application of iDTV for e-Learning.

From the information acquired in the previous stages, a model was created that presents the main relevant components for the application of iDTV in e-Learning. The proposed model is analyzed and discussed highlighting its main contributions to the theme. Finally, potentials and limitations involved in these applications are identified.

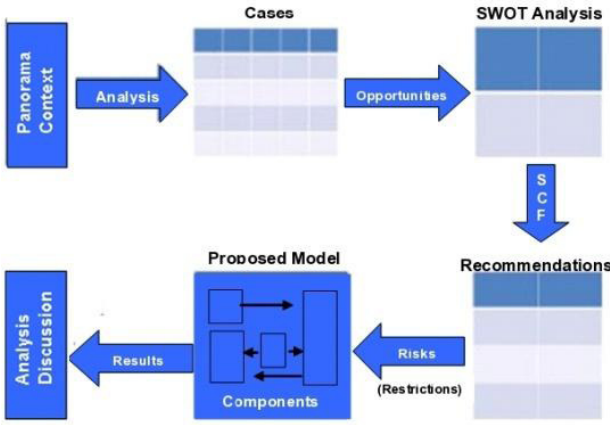


Figure 1: Methodology adopted in this study

4. E-LEARNING

According to Bates [7], even though the expression *e-Learning* is used for learning via Internet through the use of a personal computer, it can be employed to express any way of learning that uses a digital electronic device.

Its development is very dependent on the access device of the hardware/software and on the Internet provider availability. For Bates [7], although not completely reliable, content delivery is easy and demands only minimum technical knowledge to operate the computer.

T-learning reached significant importance lately, and emerges as a potential medium to create opportunities for learning at home [8]. It can be highly effective in regions where Internet access is scarce. For that matter, digital terrestrial television (DTT) is restricted to social inclusion issues. As an educational platform, interactivity via DTT is the key to reach wider audiences [8].

Bertoti [9] explains that the following factors also contribute to the use of television in e-Learning:

- *Usability:* Since it is a household appliance that has been in use for decades, the main functions of a television set (change channels, adjust volume, turn on and turn off) are already known by the population;
- *Transmission Quality:* Important media in distance education, such as audio and video, are transmitted with quality to television sets. However, that same quality will not be acceptable for e-learning applications until wideband access is universally available;
- *Information Vehicle:* unlike the computer, that can be considered a work tool, people consider the television set an appliance for information and entertainment;
- *Collaboration:* watching a TV program can be considered a social experience, in which the television program acts as mediator for interaction and collaboration among the TV viewers.

5. CURRENT APPLICATIONS OF iDTV IN E-LEARNING

Bates [7] surveyed 21 distinct applications of Digital Television in e-Learning with different interaction levels. The cases differ in terms of applicability, product, interactivity level and target audience. Distinct applicabilities were shown: service installment and information, support to educational activities, content reinforcement in some subjects, reading incentive, memorizing, and solving problems, among others. As to the final product, most of the cases have teaching and learning as their main focus. Other modalities were also present: information, entertaining, services and development of capabilities for children.

The 21 cases were classified using a simplified version developed by Crocomo [3]. Accordingly, there are three interactivity levels: "level 1", or local, does not need a return channel, and the user can navigate through several options. Interactivity "level 2" demands a return channel, allows for return messages, although not necessarily in real time. In interactivity "level 3", the return channel is mandatory, and it is possible to send and receive information in real-time. Almost all cases operate at the smallest interactive level; only 4 cases explore interaction in real-time (level "3"). Finally, regarding the target audience, the models cross over all age ranges, being specific in some cases, by varying the levels of difficulty in the proposed exercises.

5.1 SWOT Analysis

Using the implementation of BSTD-TV in Brazil as a baseline, a SWOT matrix was created, identifying the *strengths, opportunities, weaknesses and threats,* and the combination of its relations regarding *leverages, limitations, vulnerabilities and problems.* This SWOT analysis is fundamentally important for the understanding of open issues, and the submission of proposals for *t-Learning* models and services. Our evaluation is represented by the SWOT matrix in Figure 2.

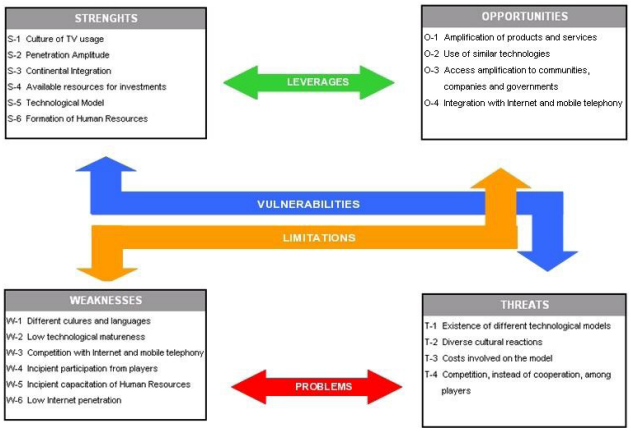


Figure 2: SWOT Matrix

5.2 Critical Success Factors

After developing the SWOT matrix applied to the scenario of the Brazilian Digital TV implementation project, we turned to the related Critical Success Factors (CSFs), in order to offer the following recommendations for the use of iDTV in e-Learning:

Table 1: Critical Success Factor and Recommendations

Critical Success Factor	Recommendations
Commitment of international and regional organisms	To involve the international regulation and development entities, in order to guarantee research financing; integration and interoperability of the technologies applied in the different global models; and the inclusion of poor communities and poor/developing countries
Development of models, frameworks and common standards	To facilitate the expansion of products, services and competitive applications with strong global penetration potential, thus reducing costs, increasing the offer portfolio and reducing market absorption time.
Mobilization of research, education and development entities	To foster the creation of models, lines of credit and the availability of specialized labor to build models, standards, products and services, as well as the generation and diffusion of iDTV contents.
Technological integration among Internet, iDTV and mobile telephony	To develop the digital convergence necessary to the application and use of different media for the process of t-Learning in a dynamic and integrated way.
Incentive to the expansion of iDTV audience in poor communities	To develop policies and models for the digital inclusion of excluded communities, through the diffusion of community usage models, equipment financing, and interactivity processes in iDTV.
Partnerships and alliances between technology and content providers	To stimulate the involvement of software, hardware, communications, education and broadcasting companies to accelerate the awareness and diffusion of offers and usage of iDTV interactive technologies by the Brazilian society.
Enhancement of public policies	To regulate the development of models, frameworks, computer applications, not to mention the protection of author rights and the use of iDTV content.

The SWOT analysis complemented with the Critical Success Factors (CSFs) and related recommendations leads to the identification and development of projects and public policies that will enable the application of t-Learning solutions using the interactivity of iDTV. Those pieces of information, along with the consideration of risks and existing restrictions in the current Brazilian model (BSTD-TV), will be the building blocks of a model for the development of iDTV applications for e-Learning, which will be presented next.

6. PROPOSED MODEL FOR THE APPLICATIONS OF iDTV IN E-LEARNING

The conception of a development model for interactive Digital Television applications in e-Learning, which meets the diverse needs imposed by the wide range of applications required in the Brazilian scenario, calls for different communication possibilities.

Even considering the low penetration of the Internet in Brazilian households, that type of access has to be contemplated in any project aimed at significant share of the population. Other media such as fixed and mobile telephony, PLC, WiFi networks and other existing networks must also be considered. Finally, we cannot forget traditional communication modes, which can be personal or through the mail, and mandatory for remote communities.

Nakayama [10] defines three essential components in an e-Learning application: the *content author*, the *content provider* and the *students*. *Content authorship* can be in charge of an organization or in charge of teachers. The *content provider* systems are, generally, provided by specialized companies or organizations. The *students* can be citizens, public or private institution employees or students of teaching networks.

The model proposed in this study, adapted from Santos [10], encompass all these diverse modes of communication that complement the iDTV application in e-Learning activities. Figure 3 represents potential interrelations between the components of the proposed model.

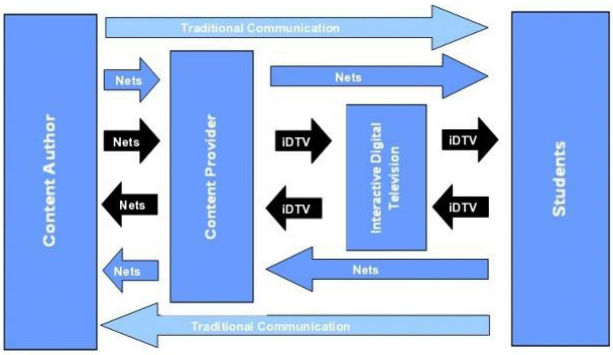


Figure 3: Proposed Model for t-Learning (adapted from [10])

iDTV's role in this model is the mediation between the content author and the students. The need for a content provider can be explained by availability and ease-of-use of authoring software , and by the flexibility of the model, including its availability through several existing local networks. In the proposed model, the content authors use network resources and the content provider to create and prepare the educational material. That material is then treated by the content provider work groups and the iDTV companies for the further diffusion of the material through open television. The model also includes the interaction between students and content providers, by means of iDTV return channels or by the Internet, or by other means of communication. The insertion of iDTV in this model increases considerably its range not only by its extensive household penetration but also by the lower operational cost for the students.

6.1 Potentials and Limitations of e-Learning Applications in iDTV

Santos [11] considers that, potentially, e-Learning will be able to reach a larger populational slice using iDTV than computers. This is due to the fact that computers are available in 12% of households, which is significantly lower than television, now

available in more than 90% of all households, according to research carried out in 2004 by IBGE – Brazilian Institute of Geography and Statistics,. Another enabling factor is that t-Learning is developed using experience acquired in years of research on e-Learning applications, due to media convergence.

The main difficulty, and the limiting factor found in the adaptation of e-Learning applications, according to [11], is that many computer input devices will not be available for iDTV Web navigators (e.g. keyboard and mouse), the size and resolution of the screen where the data will be viewed, among others.

Santos [11] also indicates that, with iDTV, the user needs to interact with the TV set, becoming an active element in the communication process, as in e-Learning programs. iDTV applications must encourage students to communicate with each other and with the teachers. Interactivity is a concept that does not exist in traditional analogical systems. In that respect, the Brazilian population is not used to interacting with television, which raises an obstacle for the successful implementation of t-Learning programs. The ideal scenario for t-Learning is one where all students have access to a return channel that provides full and free interactivity.

7. FINAL CONSIDERATIONS

In spite of the incipient use of e-Learning technologies in iDTV and the low current interactivity in the European and North American systems, as presented in the cases surveyed in this study, the BSTD-TV has, in its interactivity proposal, a very interesting potential for the evolution and development of applications and services. As a referential contribution for this process, the e-Learning models applied nowadays, if replied, offer relevant time savings in the selection, modeling and implementation of new applications and services for iDTV. One of the fundamental critical factors in that process is the equalization of models and public policies related to bidirectional interactivity by the Brazilian government.

The use of the SWOT analysis matrix enabled us to identify that the leverage of results depends on the continued investment in the development of models, standards, frameworks, applications and uses of iDTV in e-Learning, and on the increased allocation of resources for infrastructure and education of human resources. In order to minimize vulnerabilities, we proposed that the effort is focused on replying and adaptating existing models in other systems or markets, and the incentive for the development in partnerships and global consortia. As a complement to reduce the identified limitations, we propose the development of specific public policies and solutions for the digital inclusion of the poorer social layers and of specific communities.

The cases presented here confirm the technological feasibility of the application of iDTV solutions for e-Learning. The SWOT matrix and the Critical Success Factors work in an environment where those technologies are integrated with the daily life of the society. It is important that any iDTV application model in e-Learning considers the true protagonists in this process: the authors of the e-Learning process, the content provider and the students.

Further studies along these lines may investigate the application of the proposed model in a private organization or at a Higher Education institution, aiming to improve it. A research project can be conceived whereby the needs of each of the three components of the model will be evaluated, describing some technological applications that can fulfill those needs, having e-Learning as the focal point.

8. REFERENCES

[1] WINCK, João Baptista. *A promessa do audiovisual interativo*. TransInformação, Campinas, 19(3): 279-288, set./dec., 2007.

[2] ZUFFO, M. K.: *TV Digital Aberta no Brasil: Políticas Estruturais para um Modelo Nacional*. Available at: http://www.lsi.usp.br/interativos/nem/tv_digital.pdf. Accessed in August, 2008.

[3] CROCOMO, Fernando. *TV Digital e Produção Interativa: a comunidade manda notícias*. Florianópolis: Editora da UFSC, 2007.

[4] LEMOS, André L.M. *Anjos interativos e retribalização do mundo. Sobre interatividade e interfaces digitais*, 1997, Available at: <http://www.facom.ufba.br/pesq/cyber/lemos/interac.html>. Accessed in December, 2008.

[5] MONTEZ, Carlos; BECKER, Valdecir. *TV Digital Interativa: conceitos, desafios e perspectivas para o Brasil*. Florianópolis: Editora da UFSC, 2005, 201p.

[6] CPQD. *Cadeia de Valor - Sistema Brasileiro de TV Digital.* Campinas, SP, 2004. Available at: http://www.fndc.org.br/arquivos/MapeamentoCadeiadeValor-CPQD.pdf. Accessed in December, 2008.

[7] BATES, Peter J. *T-learning Study: A study into TV-based interactive learning to the home*. PJB ASSOCIATES (Inglaterra). Available at: <http://www.pjb.co.uk/t-learning/t-learning%20Final%20Report%20-%20Main%20-%205-05-03.doc>. Accessed in June, 2009.

[8] SANCIN, Chiara et al. *T-learning for social inclusion.* Available at: <http://www.elearningeuropa.info/files/media/media18506.pdf>. Accessed in June, 2009.

[9] BERTOTI, Giuliano Araujo; ALMEIDA, Felipe Afonso de; BACCAN, Davi D´andréa. *Educação à Distância Mediada pela Televisão Interativa: Panorama Atual e Perspectivas para o Brasil.* Available at: <sbie2004.ufam.edu.br/anais_cd/extras/anaisvolI/vDigital/artigos/5544.pdf>. Accessed in June, 2009.

[10] NAKAYAMA, Minoru; SANTIAGO, Rowena. Two Categories of e-Learning in Japan. Educational Technology Research and Development Journal, Vol.52, N.3, 2004.

[11] SANTOS, Davi Trindade dos. Estudos de Aplicativos de TVDi para Educação a Distância. Dissertação (Mestrado) da Universidade Estadual de Campinas - Escola de Arquitetura e Engenharia,Campinas, SP, 2007.

Conceptual Models for T-Commerce in Brazil

Bruno Cavaler Ghisi
CERTI Foundation /
Program in Computer Science – PPGCC
Department of Informatics and Statistics - INE
Federal University os Santa Catarina – UFSC
Caixa Postal 5053
88040-970, Florianópolis, SC, Brazil
+55 48 3239-2000
bcg@certi.org.br

Guilherme Figueiredo Lopes
Brava ECM
Av. Santos Dumont, 831
3802-8300, Joinville SC, Brazil
+55 47 3802-8300
guilherme.lopes@bravaecm.com.br

Frank Siqueira
Program in Computer Science – PPGCC
Department of Informatics and Statistics - INE
Federal University os Santa Catarina - UFSC
Florianópolis, Brazil
+55 48 37217748
frank@inf.ufsc.br

ABSTRACT
After creating a new standard for Digital TV, Brazil began to develop interactive content for its system. The middleware (Ginga) specification provides a solution for advanced interactivity, which allows the development of high value applications and enables solutions such as television commerce (T-Commerce).

This article proposes models to be adopted by broadcasters and advertisers for T-Commerce applications, as well as a some alternative ways of payment. There are also many difficulties in developing this type of application in Brazil. Some of them such as the secure communication norms which have not been standardized yet; integration with payment mechanisms, among others will be exposed and analyzed along this paper.

Besides analyzing these problems, this article shows the experiments in the development of two models of T-Commerce applications that could be used as a reference in the country.

Keywords
T-Commerce. Television Commerce. Interactive DTV applications. Brazilian Digital Television System. Ginga middleware.

1. INTRODUCTION
Digital television commerce, which is often referred to as T-commerce for short, is a particular case of electronic commerce using the digital television and other devices in connection with it as an infrastructure of commerce [1].

In Brazil there are still some barriers in the development of T-Commerce applications. For example, important issues related to secure communication have not been defined yet in the Brazilian Digital TV standards.

The paper presents, in Section 2, some data around digital TV and e-commerce in Brazil. Section 3 introduces the Ginga middleware and some of its characteristics that are relevant to T-Commerce. Section 4 discusses the result of a study on models for T-Commerce. Section 5 presents the experiments around the implementation of two T-Commerce applications using the proposed models. Section 6 has the conclusions and future work.

2. BRAZILIAN SCENARIO
Studies related to Digital TV in Brazil began in 2003 with the creation of the Brazilian Digital Terrestrial Television Forum (SBTVD Forum). During the lengthy discussion process about the interactivity standards, all the technological alternatives have been thoroughly studied, taking into account the characteristics of the Brazilian analog television system.

The first digital transmission was held in December 2007, in the city of São Paulo. At the moment, there are a few interactive applications for digital TV for testing purposes, without interaction from real users, because the receivers are not in accordance with the middleware standards, described in Section 3.

Television is the most common electrical appliance found in Brazilian households, present in 96.8% of them [2], largely surpassing the number of homes with computers. In parallel to this, electronic commerce (e-commerce) continues growing in Brazil, being limited only by the lack of a large computer base. In 2009, e-commerce generated a turnover of about US$ 6 billion in the country, which is 28% higher than the previous year. In the same year, a total of 17 million people had at least one shopping experience on the Internet [3].

These data demonstrate the enormous potential for the T-Commerce market in the country.

3. THE GINGA MIDDLEWARE
The middleware of the Brazilian DTV standard ISDB-TB is called Ginga [4]. Its specification is defined by the Brazilian Association of Technical Standards (ABNT). ABNT is a founding member of the International Organization for Standardization (ISO) and the body responsible for technical standardization in Brazil.

Ginga is in accordance with the ITU J.200 [5] and divides the middleware into two engines: presentation and execution. The presentation (declarative) engine is called Ginga-NCL [6] and the execution engine (procedural) is called Ginga-J [7]. There is also a bridge that allows interaction between the two environments. Finally, running above the engines, there is a common core that contains the infrastructure used by Ginga-NCL and Ginga-J. The Ginga architecture is illustrated in Figure 1.

Ginga-NCL is a logical engine that processes NCL (Nested Context Language) documents. A key component of Ginga-NCL is the declarative content decoding engine (NCL formatter). Other important modules are the XHTML-based user agent, which includes a stylesheet (CSS) and ECMAScript interpreters, and the Lua engine, which is responsible for interpreting Lua scripts [9].

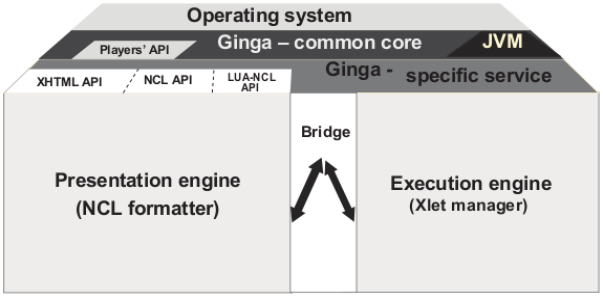


Figure 1. Ginga Architecture [6].

Ginga-J is the logical engine that handles procedural applications (Java Xlets). A key component of the procedural application environment is the procedural content execution engine, composed by a Java Virtual Machine (JVM) [9].

The bridge provides a mapping mechanism that allows two-way applications interaction on both machines, either by manipulating NCL documents in Ginga-J, executing Java code through Xlets in NCL or using the Ginga-J API in Lua scripts [6].

The common core is composed of common content decoders (such as PNG, JPEG, MPEG) and procedures to obtain contents transported in MPEG-2 transport streams and via the interactive channel [6].

Many types of applications are need connectivity back to the head-end via some form of communication route, know as return channel or interactive channel. Interactive channels come in several different flavors, depending on the cost of the receiver and the type of the network the interactive channel is connected to [10]. Ginga specifications also does not mandate the presence of an interactive channel, but in case it exists, it must be TCP/IP.

The security standard draft [11], which specifies the security requirements for Ginga, has not been concluded. The current version addresses issues related to user authentication, certificate infrastructure and permissions. It also defines that the interactive channel should support HTTP (HyperText Transfer Protocol) and HTTPS (HyperText Transfer Protocol Secure) at the application layer.

These issues around security should be completed in order to allow the development of T-Commerce applications with a level of security assurance in real environments.

4. T-COMMERCE MODELS

The literature review involving television and interactive models of T-Commerce is described in Section 4.1. The next subsections describe three characteristics that help to define models of T-Commerce: 1) presentation, 2) form of payment and 3) content associativity.

4.1 Related Work

There are not so many scientific publications related to T-Commerce, so the model proposed is based on books, commercial content and experiments.

Srivastava [12] quotes the use of T-Commerce in categories such as book, video, music, gifts, electronics, apparel and toys. Advertising is treated as something separate and he also does not relate banners to possible purchase interactions.

Gulati [13] defines three types of T-Commerce: push-commerce, enhanced shopping channels and virtual mall. It does not specify any payment models and says that the push-commerce is limited to an interactive commercial (without the possibility of doing a purchase).

Chan-Olmsted [14] considers that definitions of T-Commerce may involve purchase orders via toll-free, but much more than that, classifies T-Commerce as a) transactions conducted by clicking though with the TV remote control or b) offers presented on television and completed with an actual transaction on the PC Web Page.

4.2 Presentation Model

This article defines four conceptual presentation models for T-Commerce in Brazil that can be extended to other countries: Sales Channel, Program Related, Interactive Advertising and Other Initiatives.

Sales Channel: refers to a channel that offers a products catalog, services or an application that sells products from an Internet store. In Brazil, there are specific channels that sell products and currently do the purchasing process by telephone. For example, this model allows the possibility of having a sales channel for buying tickets for events. The sales channel could be presented during the whole day or during a given period of the program schedule.

Program Related: are applications that can appear on television programs such as interviews, football games, soap operas, among others. For example, the promotion and sale of a kind of sandal at a specific time that an actress is wearing in a soap opera or the offer of a specific book during a talk show with its author.

Interactive Advertising: adds interactivity in commercials through a specific application. The limited time of an advertising is something that need to be considered.

Other Initiatives: is related to the use of T-Commerce for other purposes. For example, donation of money for charity, buying content, gambling, among others.

4.3 Form of Payment Model

Three ways are defined as forms of payment model alternatives that can be used during the purchase process in any of presentation categories: Payment by Credit Card, Direct Debit and Other Forms. There are models that are more suitable for some applications than others, as described in Table 1. Moreover, some models can use the interactivity channel.

Payment by Credit Card: refers to the payment model where the viewer enters the credit card number and others information that are is necessary to finish the payment process. The purchase is made at the time and mandates the use of the interactive channel.

Direct Debit: the viewer only enters its user account (registered in a service) and the payment is done. Possibly it will be discounted in the next billing cycle or the payment may be generated independently. It mandates the use of the interactive channel.

Other Forms: payment is completed via any other way than television, such as telephone, email invoice, among others. The process starts on the television, but it allows sales via television receivers that do not have interactive channel. However, in the presence of an interactive channel, it could be used to provide the customer contact information to a telemarketing operator, who would make contact. It can be explored in presentation models of T-Commerce which time for application execution is limited or when an alternative to the use of the interactive channel is necessary.

Table 1. Table of possible presentation models by payment models.

Presentation Model	Payment by Credit Card	Direct Debit	Other Forms
Sales Channel	Yes	Yes	Yes ([1])
Program Related	Yes	No ([2])	Yes
Interactive Advertising	No ([3])	No	Yes
Other Initiatives	Yes	Yes	Yes

4.4 Content Associativity Model

The Ginga standard [4] presents two types of services related to TV content: contextualized and independent.

A T-Commerce application may be related to the content featured in the video. For example, the product being sold may be available for purchase at a given moment of the video. Or, an actor is wearing a watch in a soap opera and at some time, in a close-up, the application could appear making possible the purchase of the product.

Also, a T-Commerce application may not be contextualized with the content, which will be considered as the independent type. For example, a brewery sponsoring a soccer game. Table 2 summarizes the possibilities of the relationship of content in the presentation models proposed.

[1] It would be an evolution of the traditional sales channel by telephone, but adding interactivity.

[2] Usually a user does not have an account in a specific program, so it would be not possible to buy in a program in such way.

[3] There is not so much time to process a Payment by Credit Card during a commercial.

Table 2. Table of possible presentation models by content associativity.

Presentation Model	Contextualized	Independent
Sales Channel	Yes	No
Program Related	Yes	No ([4])
Interactive Advertising	Yes	No ([5])
Other Initiatives	Yes	Yes

5. EXPERIMENTAL WORK

Two model of presentation applications were developed: Sales Channel (Section 5.3) and Program Related (Section 5.4).

5.1 Assumptions

Currently there are a few implementations of the Ginga-NCL engine in virtual machines and digital TV receivers and they still do not provide the necessary resources to enable secure communication. The analysis made during this project concluded that the best way to use the security mechanisms in the Ginga-NCL would be via the Ginga-J bridge. This is because the Lua language, in Ginga, does not provide secure connections. Besides the use of HTTPS directly in NCL documents does not allow much flexibility in development.

On the other hand, in Ginga-J, the APIs to implement a secure communication have already been defined. However, there are still no implementations, either emulators or real devices, to be used in the application development.

It was also carried out a preliminary study on payment systems and third party gateways in Brazil. It was not found a widespread integration approach that meets the needs of T-Commerce. The currently available approaches allow, in most of the cases, the integration with e-commerce via browser, which does not apply to Digital TV. The final T-Commerce solution would probably need an integration via computer systems, like Web Services.

5.2 Development Process

The development was then performed on the Ginga-NCL engine through the use of Lua scripts. It was decided not to use the Ginga-J environment because of the few development tools available. The security issues described will be handled via Ginga-J bridge in future work, when the infrastructure evolves.

These application examples used an integration server (gateway) to receive purchase requests. Due to lack of integration with the payment systems at this time, it is necessary that an administrator manually complete the transaction after receiving the request. The future architecture will evolve as described in Figure 2 without manual intervention.

[4] The products being offered will be related with the program, like they do in merchandising.

[5] Interactive advertising will use interactive applications to improve the experience.

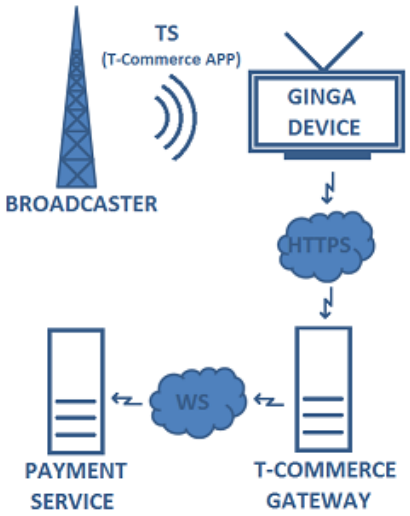


Figure 2. T-Commerce application architecture.

The applications were developed in Eclipse IDE[1]. The tests were performed locally on the virtual machine STB Virtual 0.11.2, available in Portal do Software Público[2], and also on the Proview receiver, with the RCASoft[3], a commercial implementation of Ginga middleware.

5.3 Model 1 – Sales Channel

It was developed an example of the Sales Channel category (Section 4.2), inspired in traditional e-commerce stores, such as Amazon[4].

The store is displayed in almost the whole television screen and requires full attention from the viewer, as shown in Figure 3. It offers the possibility to navigate among the various products of the store. This example have a relationship with the current channel, so it could be described as having an contextualized content associativity (Section 4.4).

Figure 3. List of products.

The viewer can also navigate and receive more details of a product (Figure 4) and add it to a shopping cart (Figure 5). This cart can be edited to change quantity or remove the contained products. All this control and interaction is done through the remote control, using color keys in the navigation actions and arrows.

Figure 4. Product details.

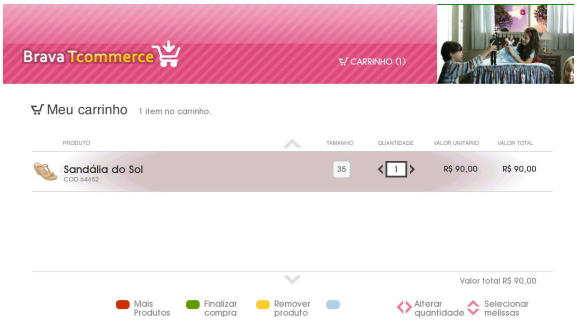


Figure 5. Shopping cart.

After checking out, a form is displayed requesting credit card information (Figure 6), following the Payment by Credit Card model (Section 4.3). After filling out the data, these are sent to a server using the JSON (JavaScript Object Notation) protocol over HTTP (Hypertext Transfer Protocol).

Once it receives the server response, the application shows a message saying if the transaction was completed successfully or not. Finally, if the interactive channel is not available in the receiver, the process of checkout displays a screen that contains a telephone, which can be used to finalize the purchase. This also demonstrates the use of Payment by Other Forms model (Section 4.3), which can be combined in this case.

Figure 6. Checkout.

[1] http://www.eclipse.org

[2] http://www.softwarepublico.org.br

[3] http://www.rcasoft.com.br

[4] http://www.amazon.com

5.4 Model 2 – Program Related

This model was developed to exemplify the category Program Related (Section 4.2). The features described in Model 1 (products list, product detail, shopping cart and checkout form) have been retained, but simplified in some way.

The application starts hidden, displaying a small non-intrusive alert at the bottom of the screen. These alerts promote certain product at an instant of the video, so it is contextualized (Section 4.4). The NCL language notifies the Lua script in certain periods of time, synchronized with the video, so the script can load the current product that will be displayed. The Figure 7 demonstrates the possibility of buying the blue clothes that the actor is wearing.

Figure 7. Alert displaying a product.

From this small alert, the viewer can navigate through more product details (Figure 8), add it to the shopping cart (Figure 9) and checkout (Figure 10). The payment model of this example is the Payment by Credit Card (Section 4.3). For each of these steps, the application takes a larger space on the screen, overlaying the video channel gradually, according to the interest shown by the viewer.

Figure 8. Product details.

Figure 9. Shopping cart.

Figure 10. Checkout.

6. CONCLUSIONS AND FUTURE WORK

This work pioneered the development of T-Commerce applications in Brazil. It proposed a categorization of conceptual models that can be exploited economically, enabling application development to promote television commerce in the country. Finally, two functional applications using Ginga have also been developed, demonstrating these concepts in practice.

A major difficulty encountered in the Ginga development was the lack of tools and testing environments. Today there are few tools to use and also the testing environments are quite immature, implementing only part of the specifications and having several bugs. It is necessary to develop such tools, as well as the environments evolve in order to improve the development of new interactive applications and take T-Commerce to the next level.

Future work includes adding security mechanisms to the applications developed – as soon as the draft is concluded - and also to use a real payment integration without requiring the use of other ways for payment.

Some qualitative usability tests are being done in order validate some ideas and improve the T-Commerce applications that were developed. More work on this subject would be very relevant to help understand the behavior of Brazilian viewers. There are just few studies in Brazil relating usability and digital TV applications experience.

Another interesting thing would be developing applications for the other proposed models and also develop the two applications described here using Ginga-J, allowing a comparative analysis. Finally, evolve these conceptual models adding other variables and exploring subcategories would be relevant.

The development of interactive applications, in Brazil, should grow from this year, when receivers with complete implementations of Ginga will be launched. Moreover, the investment done by the government, industry and broadcasters on research and development in the area of Digital TV is likely to encourage the use of interactive applications, including T-Commerce.

7. ACKNOWLEDGMENTS

This project was sponsored by a Brazilian fund for research and development called FINEP. Many thanks to all researchers involved at CERTI Foundation and Brava ECM. Finally, we would also like to thank the Graduate Program in Computer Science (PPGCC) at UFSC.

8. REFERENCES

[1] Kim, S., Lee, H. J. and Leemand, C. S. 2004. Architecture of authentication mechanism for emerging T-commerce environments. Springer Berlin, Volume 3331, Heidelberg, 540-547.

[2] Instituto Brasileiro de Geografia e Estatística. 2009. Pesquisa Nacional por Amostra de Domicílios - Síntese de Indicadores 2009. IBGE, Rio de Janeiro.

[3] e-bit E-Commerce Information. WebShoppers 20 Edition. Retrieved from: <http://ebit.com.br>. At March 2010.

[4] Associação Brasileira de Normas Técnicas. 2008. ABNT 15606-1. Televisão digital terrestre – Codificação de dados e especificações de transmissão para radiodifusão digital – Parte 1: Codificação de dados.

[5] International Telecommunication Union. 2001. ITU-T Recommendation J.200. Worldwide common core – Application environment for digital interactive television services.

[6] Associação Brasileira de Normas Técnicas. 2009. ABNT 15606-2. Digital terrestrial television – Data coding and transmission specification for digital broadcasting – Part 2: Ginga-NCL for fixed and mobile receivers – XML application language for application coding.

[7] Associação Brasileira de Normas Técnicas. 2009. 00:001.85-006/4 - Televisão digital terrestre - Codificação de dados e especificações de transmissão para transmissão digital – Parte 4: Ginga-J - Ambiente para a execução de aplicações procedurais.

[8] Soares, Luiz Fernando Gomes; Rodrigues, Rogério Ferreira; Moreno, Márcio Ferreira. Ginga-NCL: the Declarative Environment of the Brazilian Digital TV System. In Journal of the Brazilian Computer Society. No. 4, Vol. 13. p.37-46. ISSN: 0104-6500. Porto Alegre, RS, 2007.

[9] Souza Filho, Guido Lemos de; Leite, Luiz Eduardo Cunha; Batista, Carlos Eduardo Coelho Freire. Ginga-J: The Procedural Middleware for the Brazilian Digital TV System. In Journal of the Brazilian Computer Society. No. 4, Vol. 13. p.47-56. ISSN: 0104-6500. Porto Alegre, RS, 2007.

[10] Morris, S., Chaigneau, A. S. 2005. Interactive TV Standards. Focal Press, Elsevier.

[11] Associação Brasileira de Normas Técnicas. 2008. 00:001.85-005 – Televisão digital terrestre – Gestão de Direitos Digitais Parte 2 – Mecanismos de segurança opcionais. v 1.0.1.

[12] Srivastava, H. O. 2002. Interactive TV technology and markets. Artech House Publishers, 1st ed.

[13] Gulati, R., Sawhney, M. and Paoni, A. 2003. Kellogg on Technology and Innovation. Wiley, 1st ed, 144.

[14] Umar, A. 2004. Third Generation Distributed Computing Environments. Nge Solutions, 99.

An Alternative for the Interaction with Digital TV

Fábio de Jesus Lima Gomes
Computers in Education - UFRGS
Porto Alegre – RS – Brasil
fjlgomes@inf.ufrgs.br

José Valdeni de Lima
Institute of Informatics - UFRGS
Porto Alegre – RS – Brasil
valdeni@inf.ufrgs.br

Rosane Aragon de Nevado
College of Education - UFRGS
Porto Alegre – RS – Brasil
rosane@edu.ufrgs.br

ABSTRACT

This article considers the use of a barcode reader as digital TV interaction device. With the digital TV, t-learning arises as an opportunity to facilitate the learning to a large amount of people spread all over the country not covered by traditional e-learning. However, TV is not an adequate media for the reading of large texts. Alternatively, paper provides excellent readability properties and it is ubiquitous, portable, easy to use, inexpensive, can be annotated easily. But paper is a static media and does not offer capabilities such as dynamic content and linking that may be provided with digital media as a digital video in the TV, for instance. Nonetheless, the use of a remote control on digital TV requires a complex customer interaction that may cause the user's frustration. This article aims to offer an alternative linking between paper and digital TV, in order to obtain better user interface and encourage an easier interaction in t-learning.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces.

General Terms

Design, Human Factors, Measurement.

Keywords

Digital TV. Usability. Paper Interface. Learning Object.

1. INTRODUCTION

Paper is ubiquitous, highly portable, easy to use in a wide range of environments, inexpensive, can be easily annotated and provides excellent readability properties [18]. Previous research showed that the use of paper is persistent just by the physical properties of paper, not despite these properties [23]. Johnson et al. [23] claim that paper utility will not disappear with the increase of the electronic world, consequently the main goal should be integrating, not removing paper from the electronic world. Sellen and Harper [39] predict that paper, as support of reading tasks, will be one of the hardest media to be moved to the digital world.

O'Hara and Sellen [30] suggested that the critical differences (between using paper and any digital media) have to do with the major advantages that common paper offers in supporting annotation while reading, quick navigation, and flexibility of spatial layout. These, in turn, allow readers to deepen their understanding of the text, extract a sense of its structure, plan for writing, cross-refer to other documents, and interleave reading and writing. However, paper contents representation are static and do not offer any special capabilities such as dynamic content and hyperlinking that can be provided with electronic media, with a digital video in t-learning through the digital TV, for instance.

Previous research has shown that the use of remote control in digital TV produces complex interaction that can cause frustration and irritation to the user [6]. Other researchers have shown that the ease of use is one of the most important factors for digital TV [4, 9, 16, 22, 24]. In order to solve this contradiction, we present an alternative to integrating the paper with digital TV to achieve better usabilility and promote more agreeable environment to t-learning.

The next section introduces t-learning and presents advantages and disadvantages for the learning use of TV. In section 3 discusses usability emphasizing digital TV usability. Section 4 presents related work. In the section 5, a description of the developed architecture is shown. Section 6 shows the learning object implemented. Session 7 describes the usability tests. Finally, section 8 presents the conclusions and future works.

2. T-LEARNING

Bates [2] has identified a big potential for the use of iTV for increasing learning opportunities in the home, in particular through personalized options and also the need to find ways of utilizing the powerful combination of broadcast TV and interactive services to provide hooks to draw viewers into active learning environments. But iTV success requires technological solutions, sustainable models which address pedagogical issues, however, there is still limited research in the area, especially on interactivity and learning aspects [2, 25].

This is an important field in the emerging era of lifelong learning, as learning will take place in a wide variety of contexts and locations and non-formal learning will become as important as formal learning [2] demanding flexible environments. Global access to information and technology is changing the relationship between people and knowledge, and the trends in convergence, integration and co-existence of various media technologies is creating new opportunities for the globalization of learning practices.

T-learning is closely related with interactive access to video learning materials within the home or other locations like school, workplace or community learning centre, through a TV or a device more like a TV than a personal computer [2]. The same author shows a number of reasons why it is important to consider the role that interactive digital TV presents relation within a broader e-learning strategy:

- Most people have access to a television in their home;
- Not every household will have an Internet-enabled computer;
- The TV is an easy to use device;
- People tend to trust the content that is on the TV;

- The TV has the potential for reaching more people and offering learning opportunities, more than what traditional learning institutions can do.

Aarreniemi-Jokipelto [1] identified other reasons supporting the use of digital TV for learning purposes: Accessibility; Interactive services; Independence of time and place; Low threshold for starting use e Learning on demand.

Aarreniemi-Jokipelto [1] also argued that learning is an activity that should be time and place independent and digital TV supports this class of e-learning. In addition, digital TV has greater importance compared to analog broadcasting since it has two-way and feedback features. Damasio [12] identified the potential of digital TV to provide motivating, engaging, and effective media for everyone, whether the learning takes place at school, at home, or elsewhere.

3. USABILITY

Usability has been defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [21]. Usability relates to the effectiveness and efficiency of interaction among users, their tasks and the task environment [15]. Effectiveness means that the user is able to carry out the intended task, efficiency refers to the implied time, and satisfaction defines how acceptable the system is to the users [28]. Usability has been traditionally associated with and can be measured in terms of a number of attributes: learn ability, efficiency, memory ability, frequency and severity of errors, and subjective satisfaction [28]. Usability testing is always carried out to test a service or a product, not the user [28]. In learning, bad usability affects the learning process and time is wasted with the system, instead of spent in the learning process. In addition, people refuse to use a rigid, slow, and unpleasant system interface stimulating the course interruption [11].

The primary electronic multimedia device in the home is the television (TV) set with its basic input device: the infrared remote control. The transition to digital TV has changed the television concept; new services and applications are provided through this standard input interface. Recent attempts try to integrate the computer and the TV to provide a home terminal suited for entertainment as well as an information supplier [41]. These attempts have various problems, e.g. economic, technical and usability difficulties [6]. One of the main problems that challenge the home digital terminal is the complex information navigation through the remote control device. This problem is caused by the attempts to design the digital TV as a computer but with the same conceptual model of the so-called 'traditional TV' approach that produces a complex interaction [6] through remote control devices that tend to be big, containing many buttons and controls not easy to be employed [27].

New means of user interaction for the television include voice control [5, 33, 40], Personal Digital Assistants (PDAs) [36], and touchpad devices [14] which address the drawbacks of traditional remote controls such as the overabundance of buttons and controls and the lack of alphanumeric input. Another possible approach was ignored in the digital TV context is augmentation of physical paper. Previous research emphasizes the need of integration of digital computation with physical paper in order to obtain the best properties of both [37]. From an interaction standpoint, paper-based media provides many advantages that are difficult to capture in a traditional graphical user interface [39]. This approach can be a solution for the TV interaction problem.

3.1. Devices

Prata et al. apud Robertson [36] states that any attempt to create a taxonomy of devices will be a complex problem since device usage patterns change over time and vary depending on the combination being used. Thus, the best approach is to study in detail each particular situation, namely, start by studying each particular device characteristics, its specificities, and cognitive and affective aspects associated to its use; and based on that information, design the interface and the interaction model. So, we are studying the barcode reader as a digital TV device in order to design the best way to use it in the digital TV environment.

When compared with the TV, employing a PC: usually involves "one-to-one" transmission, users with different speed connections, frequent technical problems, unsafe environment due to hackers and less contents production costs; allows flexible interaction via a mouse, flexible customization, horizontal and vertical scrolling, several simultaneous windows, flexible interface, implies more homogeneous public, user individual activity and interaction, requires users lot of attention and concentration, users accept relatively delayed interactivity, users have specific goals and modes of interaction: more interactive. Compelling interface and entertainment features are desirable however not absolutely needed and ease of use [35]. When compared with a PC the digital TV implies a broadcast transmission, viewers with the same speed connection, fewer technical problems, a safer environment, expensive contents production, limited interaction via a remote control, limited customization, limited vertical scrolling, only one window at a time, a more heterogeneous public, wide audience, a relaxed and comfortable position, requires less attention; for this environment a compelling interface is fundamental, ease of use is not enough and leisure is needed [2, 10, 13, 15, 17, 34].

4. RELATED WORK

Paper as human interface has mainly been studied in the workplace environment [23, 38, 39] and compared to computer-based interfaces [20, 42]. Berglund et al. [7] have investigated the use of the paper as an interface to digital TV. They analyzed the potential of augmenting paper-based guides by linking them with TV technology. Results show that such guides help to access, read and find information. However, they are passive, isolated from the TV environment, and do not provide cognitive help. As a result, advantages and disadvantages of such guides are related to the physical properties of paper. The drawbacks may be eliminated by adding digital technology to provide interactive and TV-connected paper-based guides. The study provides design implication suggestions for a new concept called Paper Remote which is an interactive paper-based TV guide and a remote control for the TV.

Paper Remote is a concept for linking such guides with the digital TV and it allows people to access TV content with a digital pen. Results of this research indicate that digital pen and paper may be, to some extent, suitable for TV interaction. The Paper Remote concept preserves the properties of paper and at the same time adds interaction and computation technology which augment viewers' interaction. Paper Remote is a computer-augmented TV

guide that also functions as a remote control for the TV. Viewers tick designated areas on the printed guide to perform actions such as channel switching, getting more information, programming recordings of TV programs, and interacting with TV program providers. In their study an initial usability evaluation of the Paper Remote was conducted. In summary, results suggest that this concept provides straightforward interaction and is appropriate for the TV setting.

5. PAPER AS AN INTERFACE FOR DIGITAL TV

The aim of this article is to offer an alternative to support the production of educational text material printed in paper. It will also contain links for digital TV videos. These links will be printed in the form of barcode (Figure 1).

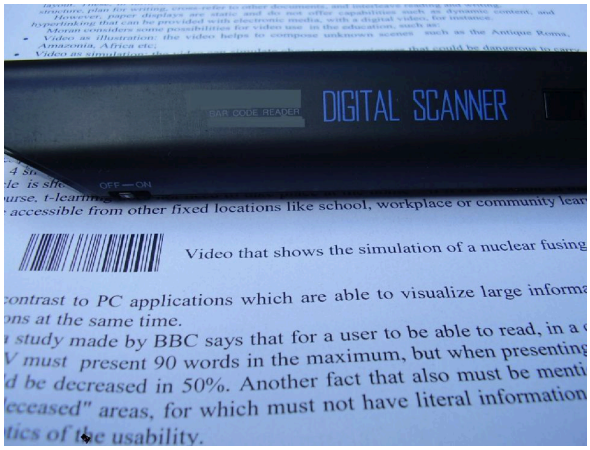


Figure 1. Example and barcode reader

Students must have the educational text material printed in paper and access to the digital TV channel offering the course, and a set-top box for download and storage of the digital videos. In the case where the student wants to watch a video about a study topic, it is enough to scan the desired link, e.g., in the barcode that is related to the video (Figure 2(1)). The barcode reader recognizes the code and sends a request to an application that is executed in the set-top box (Figure 2(2)). After that, the application locates the requested video which is stored in the set-top box and displays it (Figure 2(3).

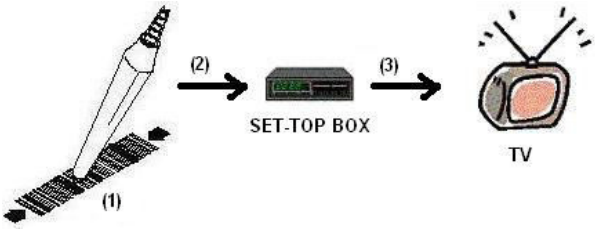


Figure 2. Example of interactive use

The interaction is local, between the user and set-top box, in the case where the user has access to a return channel, the set-top box application will send selected student information to the teacher for feedback. Digital videos are being created on the basis of the constructivist theory of Jean Piaget [32]. Among other things, this theory states that the learning occurs during the interaction of the student with the object of study. This approach applies the interactivity available in the digital TV for the creation of interactive videos turning a passive viewer into an active student.

6. INTERACTIVE SEESAW

In this paper, the above-mentioned application that it is carried out in the set-top box is a learning object (figure 3) and it can be downloaded by the data carousel and stored in set-top box. Another alternative is that the learning object may be available on the Internet and can be stored on a removable memory unit such as a pen-drive.

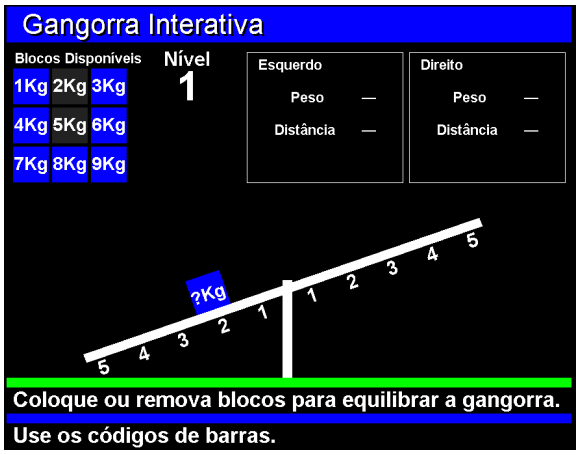


Figure 3. Interactive Seesaw

The learning object implemented is based on another learning object called Gangorra Interativa [26], which simulates a seesaw like those found in playgrounds. The objective of this learning object is to balance the seesaw, placing and / or removing blocks of different weights on the sides of the seesaw. On each side of the seesaw exists 05 (five) positions to place the weights. The weights have the following values: 1 kg, 2 kg, 3 kg, 4 kg, 5 kg, 6 kg, 7 kg, 8 kg and 9 kg.

To interact with the learning object, the viewer learner will use a sheet of paper with printed barcodes. This sheet has barcodes for the following operations: "put block on the left side", "remove block on the left side", "put block on the right side", "remove block on the right side", "next level" and "restart level".

In addition to the operations mentioned above, the sheet also has barcodes that correspond to the combination of weight with the position of the seesaw. For example, if the viewer learner wants to put the block of 6 kg in position 3 on the right side of the seesaw, he/she must first scan the barcode of the operation "put block on the right side" and then scan the barcode of the intersection weight of 6 kg at position 3. The block on the seesaw with the text "? Kg

"means that the block has its weight unknown (no information). Thus, the viewer learner must first identify the weight, so that he/she can balance the seesaw. Moreover, the viewer learner may only use the blue blocks (blocks available). The transparent blocks are unavailable.

The Interactive Seesaw has 5 (five) levels. Level 1 starts with 1 (one) block without information about its weight on the left side of the seesaw and 2 (two) blocks unavailable. Level 2 starts with 2 (two) blocks without information about their weight on the left side of the seesaw and 2 (two) blocks unavailable. Level 3 starts with 2 (two) blocks without information about their weight on the left side of the seesaw and 4 (four) blocks unavailable. Level 4 begins with 1 (one) block on the left side and 1 (one) block unavailable. In levels 1 to 4, the student may place and / or remove N blocks on both sides to balance the seesaw. Level 5 begins with 1 (one) block on the left side and 1 (one) block unavailable. At this level, the student can put only 1 (one) block on the right side to balance the seesaw.

The learning object was performed in a notebook, but it was shown on a TV screen (figure 4). To allow interaction with the object of learning, the reader barcode was attached to the notebook via a USB port. Thus, the interaction with the learning object was very close to reality as the participants did not notice that they were interacting with an application running on a notebook, but with an application running on TV. The learning object was implemented using the Java language, specifically the Java TV API. The prototype was tested with the emulator XleTView used to simulate the TV.

Figure 4. Participant in the tests

7. USABILITY TESTS

Usability tests were carried out, in accordance with the model stated by Pemberton and Griffiths [31]. The tests were performed in a room of the Federal Institute of Piauí (IFPI). Because it is a quantitative study, Nielsen [29] recommends 20 users for usability tests in this type of study. Thus, we obtained the participation of 25 students of this institute. Of the 25 participants, 6 were female and 19 were male. Only 1 (one) already had experience with DTV and 15 have used a learning object in the computer. However, none of them used a learning object in the DTV.

For the tests, the participant was placed in a room containing a sofa, a notebook, a TV and a barcode reader (Figure 4). The notebook was responsible for the storage and execution of the learning object and the transmission of audio signals and video to the TV. In tests, observation user methods, thinking aloud, interviews and usability questionnaire [28] were applied. The test was divided into three phases.

In the first phase, the participant was asked to fill in the consent form and a pre-test questionnaire with some personal information (name, age, sex etc.) and information about him/her experience in the use of DTV and learning objects. At this stage, the material printed with barcodes was shown to participants and demonstrated how to use the reader to scan the barcode. In addition, it was shown how to use the learning object. At this stage, we sought to obtain the participants familiarity with the reader barcode. In the second phase, participants had a few minutes to handle the reader barcode, but also to use the learning object.

In the third phase, was asked the participant to initiate the test. At this point, the participant should use the barcode reader to use the learning object. At the end of the test, the participant was interviewed and he/she was asked to fill out a questionnaire with questions about the use of the learning object and the SUS usability questionnaire [8]. Each session lasted around 26 minutes.

In the tests, errors were not detected by the participants in the act of scanning the barcodes. The SUS questionnaire results were also quite favorable, achieving a score of 78.4 points on a scale of 0 to 100 points. The test results showed that there is the possibility of using the barcode reader like a device to interact with the TV.

8. CONCLUSIONS

The aim of the research presented in this paper is to propose the use of a reader barcode as a device to interact with digital TV (DTV), using an interface of paper with printed barcodes. As contributions, we can highlight:

- The linking of media (paper and TV) in order to promote better usability for DTV, through a technologically simple and relatively inexpensive;
- The design of a new remote control for the DTV, with a reader barcode embedded. It is not intended to replace the remote control, but offer an alternative interaction with the TVD using a barcode reader.
- The potential of the learning object to support the conduct and implementation of cognitive strategies for problem solving to allow the construction of knowledge.

For the students who participated in the tests, there was awareness of the pedagogical value of the learning object. The test results lead to the conclusion that the technology was considered friendly and that the problems were considered normal, part of a process of taking note of a new technology. This notion is fundamental to the successful implementation of a new media or new technology..

It is evident, as described in the previous sections, the importance of the integration of paper as textual content distribution and the digital TV for the education. The main advantage of such media integration is to make the digital TV more user friendly. As a consequence of the better usability, it is expected a major increase in the motivation and learning.

As future works, it is intended to develop a model for the educational text material printed in common paper and electronic material, and to build an authoring tool for the material printed and digital videos.

This work is integrated with a national research oriented to support the Brazilian digital public TV implementation and will contribute to the user centered design and its results will influence the interaction design with digital TV.

REFERENCES

[1] Aarreniemi-Jokipelto, P. (2006). *Modeling and Content Production of Distance Learning Concept for Interactive Digital Television*. PhD thesis. Department of Computer Science and Engineering, Helsinki University of Technology, Finland.

[2] Bates, Peter J. (2003). *A study into TV-based interactive learning to the home*. pjb Associates, UK. 2003.

[3] BBCi. (2002) *Interactive Television Style Guide*, version 2.1, 2002.

[4] Black, A.; Bayley, O.; Burns, C.; Kuuluvainen, I.; Stoddard, J. (1994). *Keeping viewers in the picture: real-world usability procedures in the development of a television control interface*. In Conference companion on Human factors in computing systems, pages 243 – 244, Boston, Massachusetts, United States. ACM Press.

[5] Berglund, A.; Qvarfordt, P. (2003) *Error Resolution Strategies for Interactive Television Speech Interfaces*. In Proceedings of the 9th IFIP TC13 International Conference on Human-Computer Interaction*, INTERACT 2003* (September 1-5, Zurich, Switzerland), Swiss Federal Institute of Technology (ETH).

[6] Berglund, A. (2003). *Augmenting the Remote Control: Studies in Complex Information Navigation for Digital TV*. PhD thesis. Department of Computer and Information Science. Linköpings Universitet, Sweden.

[7] Berglund, A; Berglund, E.; Larsson, A.; Bang, M. (2005). *The Paper Remote: An Augmented TV Guide and Remote Control*. International Journal: Universal Access in the Information Society (UAIS), Springer-Verlag Heidelberg.

[8] Brooke, J. (1996). SUS - *A quick and dirty usability scale*. Available in http://www.usabilitynet.org/trump/documents/Suschapt.doc Accessed on December 02, 2009.

[9] Choi, H.; Choi, M.; Kim, J.; Yu, H. (2003*). An empirical study on the adoption of information appliances with a focus on interactive tv*. Telematics and Infomatics, 20(2):161–183.

[10] Chorianopoulos, K. (2003). *The Virtual Channel Model for Personalized Television*. *In Proceedings of the EuroiTV2003.*

[11] Costabile, M., De Marsico, M., Lanzilotti, R., Plantamura, V., Roselli, T. (2005) *On the Usability Evaluation of E-Learning Applications*. Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 3-6 Jan. 2005.

[12] Damásio, M. J. (2003). *Uses of Interactive Television on Educational Settings: Evaluating the Media Impact*. In Proceedings of the European Conference on Interactive Television, EuroITV2003, Brighton, United Kingdom.

[13] Dimitrova, N. et al. (2003). *Content Augmentation Aspects of Personalized Entertainment Experience*. *In Proceedings of the Workshop on Personalization in Future TV, UM'03* (42-51).

[14] Enns, N. R .N.; MacKenzie, I. S. (1998). *Touchpad-Based Remote Control Devices*. In Proceedings of the CHI 98, Conference on Human Factors in Computing Systems. Los Angeles, USA, pages 229-230. ACM Press.

[15] Eronen, L. (2004). *User Centered Design of New and Novel Products: Case Digital Television*. PhD thesis. Telecommunications Software and Multimedia Laboratory, Department of Computer Science and Engineering, Helsinki University of Technology.

[16] Freeman, J.; Lessiter, J. (2003). *Using attitude based segmentation to better understand viewer's usability issues with digital and interactive TV*. In Proceedings of the 1st European Conference on Interactive Television: from Viewers to Actors ? (EuroITV), pages 19-27. Brighton, United Kingdom

[17] Gill, J. & Perera, S. (2003). *Accessible Universal Design of Interactive Digital Television*. *In Proceedings of the EuroiTV'03*, UK, Brighton.

[18] Hansen, W. J.; Haas, C. (1988). *Reading and writing with computers: a framework for explaining differences in performance*. Communication ACM, 31(9): pages 1080-1089.

[19] Harper, R. H. R.; Hughes, J. A.; Shapiro, D. Z. (1991). *Working in harmony: An examination of computer technology in air traffic control*. In Bowers, J. M. and S. D. Benford, editors, Studies in Computer Supported Cooperative Work: Theory, Practice and Design, pages 225-234. Amsterdam: Elsevier.

[20] Hornbek, K.; Frokjer, E. (2003). *Reading patterns and usability in visualizations of electronic documents*. ACM Transactions on Computer-Human Interaction, 10(2): pages 119-149.

[21] ISO 9241-11. (1998) *Ergonomics requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability*. International Standard.

[22] Jan, V. D.; Loes, D. V. (2001). *Searching for the holy grail: Images of interactive television*. new media & society journal, 3(4):443 – 465.

[23] Johnson, W.; Jellinek, H.; Leigh K. J.; Rao, R.; Card, S. K. (1993). *Bridging the paper and electronic worlds: the paper user interface*. In Proceedings of the SIGCHI, Conference on Human Factors in Computing Systems. Amsterdam, The Netherlands, pages 507-512. ACM Press.

[24] Kang, M.-H. (2002). *Interactivity in Television: Use and Impact of an Interactive Program Guide*. Journal of Broadcasting and Electronic Media, 46(3): pages 330-345.

[25] Lytras, M.; Lougos, C.; Chizos, P.; Pouloudi, A. (2002) *Interactive Television and e-Learning Convergence: Examining the Potential of t-Learning*. *European Conference on e-Learning 2002*, pages 1-11.

[26] Melo, B. R. S.; Vasconcelos, F. H. L.; Pequeno, M. C.; Castro Filho, J. A.; Silva, V. M. L. *Objeto de Aprendizagem Gangorra Interativa na Compreensão Conceitual de Grandezas*. Anais do XXVIII do Congresso da SBC (CSBC 2008), WIE – Workshop sobre Informática na Escola. Belém, PA, Brasil

[27] Nelson, M. (1999). *Remote controls*. DigiPoints: The Digital Knowledge Handbook, 3(4).

[28] Nielsen, J. (1993) *Usability Engineering*. AP Professional, Cambridge, MA, SA.

[29] Nielsen, J. (2006). *Quantitative Studies: How Many Users to Test?* Available in <http://www.useit.com/alertbox/quantitative_testing.html>. Accessed on December 02, 2009.

[30] O'Hara, K.; Sellen, A. (1997). *A Comparison of Reading Paper and On-Line Documents*. In Proceedings of CHI 97, Atlanta, USA, pages 335-342. ACM Press.

[31] Pemberton, L.; Griffiths, R. N. (2003). *Usability Evaluation Techniques for Interactive Television*. In Proceedings of the HCI International Conference. Crete, Greece, pages 882-886.

[32] Piaget, Jean. (1990). *Para Onde vai a Educação ?* Livros Horizonte, LDA. 1990.

[33] Portolan, N.; Nael, M.; Renoullin, J-L.; Naudin, S. (1999) *Will We Speak to Our TV Remote Control in the Future ?* In Proceedings of the 17th International Symposium on Human Factors in Telecommunication (HFT 99). Copenhagen, Denmark. Information Gatekeepers, Inc..

[34] Prata, A. (2005). An *Overview of iTV Guidelines - A New and Critical Research Area*. In *Encyclopedia of Multimedia Technology and Networking*.

[35] Prata, A.; Chambel, T.; Guimarães, N. (2007) *e-iTV: Cross-Media Personalized Learning Environments via Interactive TV*. In Proceedings of the 5th European Conference on Interactive Television (EuroITV 2007) pages 107-113.

[36] Robertson, S.; Wharton, C.; Ashworth, C.; Franzke, M. (1996) *Dual Device User Interface Design: PDAs and Interactive Television*. In Proceedings of the CHI 96 Conference on Human Factors in Computing Systems. Vancouver, Canada, pages 79-86. ACM Press.

[37] Scaife M.; Rogers, Y.; Halloran, J. (2001) *eSpace: Integrating Novel Displays and Devices for Augmenting Collaborative Transactions*.

[38] Sellen, A.; Harper, R. (1997). *Paper as an analytic resource for the design of new technologies*. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Atlanta, Georgia, United States, pages 319-326. ACM Press.

[39] Sellen, A. J.; Harper, R. H. R. (2001). *The Myth of the Paperless Office*. Cambridge (MA): MIT Press.

[40] Tan, G.; Takechi, M.; Brave, S.; Nass, C. (2003) *Effects of Voice vs. Remote on U.S. and Japanese User Satisfaction with Interactive HDTV Systems*. In Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems. Ft. Lauderdale, FL, USA, pages 714-715. ACM Press.

[41] Voutsas, D.; Halverson, C. (2000). *Surfing the home with your TV*. In Proceedings of the 8th ACM International Conference on Multimedia, pages 452-455. ACM Press.

[42] Iatrino, A; Modeo, S. (2006). *Text editing in digital terrestrial television: a comparison of three interfaces*. In Proceedings of the 4th European Conference on Interactive Television: Beyond Usability, Broadcast, and TV. (EuroITV 2006), pages 198-204.

The Importance of Digital TV for Countries in Development: a case study of Brazil

Hildegard Paulino Barbosa
MUS3-UFPB
hildegardpaulino@gmail.com

Julio Cesar Ferreira da Silva
LAVID-UFPB
Cidade Universitária - João Pessoa - PB - Brasil - CEP - 58059-900
julio@lavid.ufpb.br

Ed Porto Bezerra
UFPB
edporto@di.ufpb.br

ABSTRACT

The Digital TV deployment allows to think a new paradigm for the vehicles of mass communication, by inserting mechanisms of interactivity. This means a new way to transmit content and connect the people of a nation and, thus, eventually shows up on various social aspects of a country. Brazil is undergoing a period of definitions and experimentations regarding the implementation of these mechanisms in the programs for Digital TV. Therefore, many aspects (technicians and mainly socials) must be taken into account in this phase, because due to the interactivity, it is expected that there will be changes in education, industry, social inclusion, among other areas.

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Sciences

General Terms

Economics, Human Factors

Keywords

TV Digital, Society, Education, Economy, T-Commerce

1. INTRODUCTION

The television, like other means of mass communication, follows the global trend of digitalization through an accelerated process of replacing its analog plataforms to digital platforms interoperable.

The impacts of the digitalization won't be concentrated only on the technologies involved in the encoding, transmission and capture, but also in various aspects of society such as education, economy, trade, governance and others. The amount of change in these aspects can be greater than expected and, therefore, caution should be exercised in decisions to be taken and how the process will be taken by the authorities.

Another major impact to be felt after the adoption of the standard Digital Terrestrial TV (DTT) is the need to develop new business models that: (i) encourage people to invest in equipment of new generation TV and (ii) allow the broadcasting stations obtain return on investments made. According to this view, the resources used by television networks and audience must be seen as investments for the exploitation of new business. The highest expected return is the interaction.

This paper discusses the importance of digital TV for the Brazil, the changes and the risks that it brings for the country.

2. THE IMPORTANCE OF THE TV IN BRAZIL

Studies done by the Instituto Brasileiro de Geografia e Estatística (IBGE) shows that the percentage of households with television has increased since 1992, when the research began to be made. In that year, the percentage was 74% while in 2008, 96.8% [1]. These percentages represent ownership, in Brazilian residences, of the receivers of analog signal, the only in use until December 2, 2007. Since then, the broadcasting stations started to broadcast digital signal.

Thus it is possible to verify that, in a country where the social inequality and exclusion of goods and services is present in all regions, there is one means of communication that unites all classes indistinctly and where any information transmitted achieves the overwhelming majority of its huge population.

The possibility to offer services through a means of communication with this reach makes it possible to think in the eradication of an attribute unfortunately present in almost all the southern hemisphere: the exclusion of services. The unidirectional transmission (sense station-viewer), inherent in analog format, makes this impossible, but with the interactivity, innovation of Brazilian System of Digital Television (BSDTV), it is possible to think in ways to reduce this exclusion and so to make a society more egalitarian.

2.1 The Digital TV

A system of Interactive Digital TV (DTV) is a consequence of digital TV. One of biggest differentials of the signal transmission of digital television will be interactivity. With the DTV, an user leaves of be just a receiver of the signal and can also to send data of your interest, such an e-mail, requisition a web page, current account balance, response to a survey or files via torrent etc. According to [4], to send data to other people, the viewer answers a need inherent of the human being, which is to participate, give their opinion and feel more embedded in their social context.

2.2 BSDTV goals

The three main BSDTV objectives are resumed in follow.

- Promoting social inclusion, cultural diversity of the country and native language through access to digital technology, aimed at the democratization of information;

- Propitiate the creation of universal network of distance education;

- Stimulate research and development and propitiate the expansion of Brazilian technologies and national industry related to information technology and communication.

The BSDTV, perhaps unlike other digital TV systems in use today, has broader objectives and more focused on the social aspects. Among them, there is the digital and social inclusion and aid to distance education. This means that the deployment of the technology digital transmission in Brazil doesn't only aims modernize the system of transmission, but also to facilitate access of Brazilians, mainly the poorest, to digital technology.

3. CHALLENGES

The five main BSDTV challenges are education, economy, T-gov, T-commerce and T-banking. Each of these areas is discussed in follow.

3.1 Education

The current scenario of Brazilian education is very serious. In 2008, 10.4% of Brazilians were considered illiterate [2], which represents more than 19 million people, according to an estimate made by IBGE in July 2009 [3]. The public schools have a much lower quality than the private schools, which means that the poorest have no access to a quality education. In addition, teachers who teach in these institutions are sometimes not sufficiently trained for the profession. The teaching model is considered by many experts (and even by the Ministry of Education and Culture of Brazil) as overshot [4, 5]. The student performance, even in the best schools in the country is not always the most satisfying.

A major problem that causes the students do not have their best performance is the lack of encouragement for their studies. Thanks to interactivity, digital TV could circumvent this problem. Combining the ease of manipulation of the apparatus with the possibilities of the digital technology embedded, a student might feel more encouraged to learn. Similarly a teacher could pass contents of a much more exciting way. A good suggestion for use of this technology is the development of applications for digital TV directed on education. These applications could be questions and answers about a particular subject in which student performance would be measured by the number of correct answers.

The use of digital TV for education is far more palpable than the computer, due to the price of the unit (it is assumed here that after a while the price of digital TV will be equal to the price of the current analog TVs, which are clearly cheaper than many of the personal computers). The computer stimulates individual use - one computer per student -, unlike television, which can be used collectively so comfortable. As the equipments and the applications are cheap (the latter could even be free, according to the ideology of free software), this new teaching tool may reach the majority of the population. The television, which currently can be seen as a problem to education, will become an ally in the teaching of students in Brazil, thanks to the interactivity and its ability to connect to the internet. Thus, it is created a new term: the t-learning, i.e. learning through digital television, similar to e-learning, which almost always refers to learning via Internet.

Television can be used not only as an aid for alive teaching, but also, or rather, mainly to distance education. As mentioned in the section 2.2, among the goals that the Brazilian government has planned for the introduction of digital TV is the creation of a universal network of distance education. The intention is to allow students, wherever they are, can communicate with a teacher in order to ask questions, discuss or learn about a certain topic.

3.2 Economy

The digital TV may also be an opportunity for the Brazilian electronics industry. There are currently some efforts to strengthen the hardware industry in the country, such as the Brazil-IP (Brazil Intellectual Property), which aims to build a Brazilian market for integrated circuits [6]. By the fact Brazilian digital television to not be a imported technology, it could change the sad reality of the country (as regards the production of hardware components). Nowadays, it practically does not add intellectual value to any of the products of this industry sold here, that is, it only import them and, perhaps, mount them [7].

In addition to reducing the volume of imports reached today, Brazil has now the chance to export technology. Among the three digital TV systems in use today, only the Japanese is used only by him. The U.S system was adopted in Canada and South Korea, while the European is also used in Australia, India, Malaysia and South Africa [8]. As the Brazilian system is the most modern, it may also be adopted in other countries. Venezuela, Peru, Argentina and Chile adopted the Brazilian Digital TV, due to the possibility of transmission to mobile devices with no additional cost. The Brazilian government is still negotiating with Cuba and Ecuador.

3.3 T-gov

The T-gov, or government via television, is derived from the E-government (electronic government). The difference between the T-gov and E-gov is the medium of access that in the first case is the digital television. It consists in to ease access to government data and public services of interest to society, besides to make more efficient internal operations of government. This means that it makes possible the population have access to the accounts of the government of your city, may examine it and denounce it by any irregularities. In the case of Brazil to make appointments to the Healthy Universal System, to get passports, enroll yourself in an exam for a university or a governmental post etc. just by using a computer with internet access.

The problem with this approach is precisely the lack of access to it by the population. The great advantages of this new form of government have little utility if the population has no access to the means necessary to obtain them. Only 23.8% of Brazilian households (representing more than 13.7 million households) have access to the internet [9], essential means for the realization of electronic government. This means that 45.21 million people (considering the average number of residents per household of 3.3) must dislocate to government offices, face long queues if they want to solve some pending subject with the government, either

by lack of conditions to obtain a computer with internet access or just the inability to use the appliance. Very clearly this new format is good, but currently is not for everyone.

Taking advantage of the reach that the analog TV has in the country and of the possibilities that digital technology offers, combined with the ease of use of television, the country may gain even more efficiency in government. If digital television get the same presence than analog TV currently have, the poorest people will achieve most of the services to the government without leaving home. The efficiency would not be only within government offices, but also on the internet, on tasks of interest and accessible to (almost) the whole society.

3.4 T-commerce

This new technology not only promises to increase foreign trade, but also the internal. Thanks to interactivity, a new type of trade may arise in the country: the T-commerce, i.e. trading on television. According shows [10], in Japan is possible to buy any product through the digital TV (or mobile devices that receive the signal) as soon as it appears on the screen. It will also be possible in the Brazil.

Any reason that leads a person to decide not buying a product in the traditional commerce or the internet (i.e. forgetfulness or insecurity) may be eliminated with the trade on television, because it will always be available to the viewer while he watches the programs, beside to promise sale quickly and easily through the device. This results in a higher volume of sales and all the good consequences of that fact. Only in Brazil, the e-commerce turnover of R$ 4.8 billion in the first half of 2009 [11], showing continued growth. The T-commerce, enjoying this growth and the increased public confidence in this type of service, may move similar or larger amount of money, given the supposed ease of it and also because it is always "in the right place at the right time."

3.5 T-banking

T-banking is the manipulation of accounts of an user via television. The idea is that the bank's customer can do everything do on the internet (or even in person), but through the television, including balance inquiry, payments, transfers and other services. This new means of access to the bank promises to increase the number of transactions made at a distance, to facilitate remote access to the bank and to gain new customers, mainly due to number of people that uses television and the facility that the applications developed to the new platform should offer. One of the first applications of T-Banking belongs to the official bank, called Caixa Econômica Federal. It provides funding of real properties by the digital TV. It is an application made by NCL/Lua [12] that displays the amount financed, the value of the installments, the interest rate etc..

4. RISKS

The two main BSDTV risks are in T-commerce and increase of the digitally excluded. They are resumed in follow.

4.1 T-commerce

The heating of the internal market through the T-commerce promises to increase the internal volume of sales due to the easy reach, speed and ease in shopping by television, attributes not always present in other types of trade. The problem is the potential indebtedness in the new business format. In [13] shows that this debt reached 34.8% of the annual income in June 2009, an increase of 8.1 percentage points compared with the previous two years, which shows the power of many Brazilians to run into debt. At the same time, [14] reported that the default rate in November 2009 reached 5.8% for credit. For individuals, this percentage is 8.1%. The T-commerce may increase this rates, due to the amenities of buying of the format and to (ugly) habit of some people do not pay off all your debts or spend more than have (especially the so-called compulsive consumers). This may result in the elevation of the quantity of dirty names in services protection credit or even, in one extreme case, in an economic crisis, if the default rate becomes too high.

4.2 Digital Exclusion

The first and major risk is the possibility of the poorest part of society to be left out of this process, at least initially. Although has been mentioned that digital TV likely will to include digitally this layer (even as because a goal is guaranteed by law), must be taken into account that they will only have the equipment and will make full use of all its possibilities and features if they know handle it. Even promising a powerful and simple interface, which will to make possible the use even by people that little (or nothing) handled digital applications in their life, these people should be educated about all that digital TV offers them. Moreover, they also need to know the meaning of the terms most commonly referred to in the transmission and the terms that they need to understand while handling. Otherwise, the Brazilian television, the only means which binds the layers together without distinction, will be another that differentiates people also by their class and level of education. In addition, one of the explicit objectives of the introduction of digital TV, which is the social and digital inclusion, will be increasingly far from being achieved, so that much investment will, in theory, "in vain".

Another risk is the price of set-top boxes. There are announcements of efforts to lower the price of the converter [15], but if no effective measure is taken, again the poorest Brazilian population will be likely to stay out of the process of modernization of its television system. The Ministry of the Casa Civil aims to implement a plan to leave the set-top boxes to levels between R$ 100,00 and R$ 120,00, but 42 million people live on minimum wage (R$ 510,00), either in actual wages or social security benefits [16]. That is, even with the relatively low price, tens of millions of people will have to spend 1/5 of their income if they want receive the digital signal. This means that the only way that will allow people to buy these converters is the little installments. And to top it off, there is a growing trend of converters is not sold separately [17]. According to Marcelo Martins, director of new business of Century (a Brazilian company that operates in the field of satellite dishes), the manufacturers decided that because the converters sold separately lack of profits. If this is confirmed, all must bear the costs of purchasing the converter more the television, even those who live on minimum wage.

5. CONCLUSION

It is valid to conclude that digital TV is promising to Brazil and has potential to develop the country in many ways: economic, social, technological, educational and others.

The decision to create the system itself was the right decision - although perhaps the decision to import the Japanese system allowed the full operation nowadays - and can lead the country to be a relatively important exporter of technology, especially in South America, "status" that it is getting just recently.

In addition, this technology can be used as a means of distance education, that is, has the potential to be an effective and efficient mean of distance education, a goal that is provided by law. And certainly a country that invests in education will progress in future.

The progresses that it can bring to the economy, governance and commerce of the country are also promising and may, in extreme cases, cause a profound impact on quality of life of the population.

However, it is very unlikely that all the benefits of this deployment are achieved. There are various "hidden" forces in the country that are quite comfortable with the position they occupy now in relation to the rest of the population and will act aiming to prevent any great transformation. Throughout the Brazil's history has been verified that these forces tend to be successful in their attacks. So that even with all this potential of development, there is a serious risk of the country don't develop and all the evils of modern society remain intact in future times.

6. ACKNOWLEDGMENTS

Our thanks to CAPES, the entity of the Brazilian Government for training of human resources that supported this work.

7. REFERENCES

[1] IBGE. 2009. Síntese de indicadores sociais: uma análise das condições de vida. DOI = http://www.ibge.gov.br/home/estatistica/populacao/condicaodevida/indicadoresminimos/sinteseindicsociais2009/default.shtm

[2] BBC BRASIL. IDH do Brasil cresce puxado por taxa de alfabetização. O Globo, Rio de Janeiro, 18 dezembro 2008, Mundo. Disponível em: <http://oglobo.globo.com/mundo/mat/2008/12/18/idh-do-brasil-cresce-puxado-por-taxa-de-alfabetizacao-587355672.asp>. Accessed December 17, 2008.

[3] G1. População brasileira ultrapassa os 191 milhões, diz IBGE. G1, São Paulo, 14 agosto 2009, Brasil/IBGE. Disponível em: <http://g1.globo.com/Noticias/Brasil/0,,MUL1266903-5598,00-POPULACAO+BRASILEIRA+ULTRAPASSA+OS+MILHOES+DIZ+IBGE.html>. Accessed December 17, 2009.

[4] LOTURCO, Bruno. Modelo educacional errado ameaça democracia. Universia, 25 junho 2009. Disponível em: <http://www.universia.com.br/ead/materia.jsp?materia=17864>. Accessed December 17, 2009.

[5] AGÊNCIA ESTADO. MEC quer ampliar carga horária do ensino médio. Último Segundo, 5 maio 2009. Disponível em: <http://ultimosegundo.ig.com.br/brasil/2009/05/05/mec+quer+ampliar+carga+horaria+do+ensino+medio+5938929.html>. Accessed December 17, 2009.

[6] BRAZIL IP. Brazil IP Network. Disponível em: <http://www.brazilip.org.br/>. Accessed December 17, 2009.

[7] ZUFFO, Marcelo Knörich. TV digital aberta no Brasil – Políticas estruturais para um modelo nacional. São Paulo. Disponível em: <http://www.lsi.usp.br/interativos/nem/tv_digital.pdf>. Accessed December 17, 2009.

[8] ESPECIAL: TV DIGITAL. Brasíia: TV Câmara. Disponível em <http://www.youtube.com/watch?v=Oo8xIiND3xU&feature=PlayList&p=C2F5853141072B65&index=23>. Accessed December 17, 2009.

[9] IBGE. Pnad 2008: Mercado de trabalho avança, rendimento mantém-se em alta, e mais domicílios têm computador com acesso à Internet. Disponível em: <http://www.ibge.gov.br/home/presidencia/noticias/noticia_visualiza.php?id_noticia=1455&id_pagina=1>. Accessed December 17, 2009.

[10] TV Digital revoluciona maneira de assistir televisão. César Menezes e Roberto Kovalick. São Paulo: Rede Globo, 2009. (9 min): Digital, color. Com narrativa. Jornalístico.

[11] Brazilian e-commerce grows 27% in the first half of 2009. http://idgnow.uol.com.br/internet/2009/08/18/comercio-eletronico-brasileiro-cresce-27-no-primeiro-semestre/, Accessed December 18, 2009.

[12] CRUZ, Daniel A. da, O que é a TV Digital?. Associação Brasileira de Emissoras de Rádio e Televisão. Disponível em < www.abert.org.br/novosite/biblioteca/TVDigital.pdf >. Accessed in February, 7, 2010.

[13] OLIVEIRA, Kelly. Relatório aponta aumento do nível de endividamento das famílias. Agência Brasil, 26 junho 2009. Disponível em: <http://www.agenciabrasil.gov.br/noticias/2009/06/26/materia.2009-06-26.3551249444/view>. Accessed in December 18, 2009.

[14] AGENCIA ESTADO. Inadimplência média fica em 5,8% em outubro. Época Negócios, 25 novembro 2009. Available at: <http://epocanegocios.globo.com/Revista/Common/0,,EMI106512-16418,00-INADIMPLENCIA+MEDIA+FICA+EM+EM+OUTUBRO.html>. Accessed December 18, 2009.

[15] ZMOGINSKI, Felipe. Conversor custará até R$ 120, diz Casa Civil. Info Plantão, 19 outubro 2009. Disponível em: <http://info.abril.com.br/noticias/tecnologia-pessoal/conversor-custara-ate-r-120-diz-casa-civil-19102009-48.shl>. Accessed December 17, 2009.

[16] OLIVEIRA, Nelson. MP eleva salário mínimo para R$ 465. Agência Senado, 27 maio 2009. Disponível em: <http://www.senado.gov.br/agencia/verNoticia.aspx?codNoticia=91542&codAplicativo=2>. Accessed December 17, 2009.

[17] DOS SANTOS, Alex. Conversores para TV digital: rumo à extinção?. Planet Tech, 2 setembro 2009. Disponível em: <http://planetech.uol.com.br/?cat=33>. Accessed December 17, 2009.

An Analysis of iDTV Appropriation in Developing Countries from the Users' Behavior Involved in Community-Oriented Content Creation Process

Elizabeth S. Furtado
UNIFOR/UECE
Washington Soares,1321
Fortaleza, Brasil
Elizabet@unifor.br

Vasco Furtado
UNIFOR/SETICE
Washington Soares,1321
Fortaleza, Brasil
vasco@unifor.br

Edson Santiago
UNIFOR
Washington Soares,1321
Fortaleza, Brasil
edsonsantiago@uol.com.br

ABSTRACT
The arrival of Digital Television in Latin America has been preceded by a huge expectation that—through its interactive resources—it will alleviate problems related mainly to a lack of information, communication, and limited entertainment options. This paper describes an iDTV experience with users in Brazil during the realization of a European Project. Supporting the interactivity and the content production was the main focus of this experience. The results indicated the main challenges for the iDTV development. Developing Country's economic growth and cultural aspects of people can help researchers to find alternatives to enrich the users' iDTV experiences regarding transition between iDTV and other media.

Categories and Subject Descriptors
D H.1.1 [models and principles]: Systems and Information Theory

General Terms
Design, Human Factors, cross-midia transition

Keywords
Multi-national driven projects, cross-cultural stakeholders, economic growth, community oriented content creation.

1. INTRODUCTION
The Terrestrial Brazilian Digital TV System (SBTVD) was launched in December 2007, since then, 34 cities —most of them state capitals—now have digital broadcast TV (but it only includes high definition transmission with no case of interactivity). The established deadline to access this technology in all Brazilian municipalities (n=5600) is 2016. Some factors of this system (e.g. existence of programming languages, declarative and imperative to develop interactive applications for Digital TV (iDTV)) will require more content to be conveyed through these applications. The production of content that exploits the interactivity offered by the SBTVD is still restricted. There have been over the past two years, prototypes of instructional applications, often encouraged by the Brazilian Forum for Digital Television, an organization that brings together universities, media companies and government agencies. As result of this support, we highlight the production of content for distance learning via the digital TV promoted by universities. It is worth mentioning that they are experimental programs, most in a testing phase. Universities are considered herein as spaces of TV content production.

Exploring this idea of content production, Mayora and his friends [1] mentioned in their paper there is a new tendency and expectation —mainly from iDTV developers´ point of view—: "The transition from mere content consumers to more proactive prosumers that transformed the web is finding now place in the context of iDTV. In the close future it will be common that TV content will be empowered with interactive applications that will be created not only by professionals but also on the final users' side. This fact will change the traditional role and stereotype of TV users from "coach-potatoes" to active generators of content".

Motivated by this tendency and considering also communities built around TV can be naturally associated to specific local services; we developed the SAMBA project [2], from 2007 to 2009. In this project, the objective was to provide rural communities with means for creating and consuming relevant content, mostly impacting on digital inclusion for the target region. Specifically, SAMBA aimed at providing specific communities and territories (such as rural areas affected with low accessibility to the Internet in developing countries) with interactive services focusing on the habitants´ local needs.

In this paper we focus in one of the three main activities, which composed the project [3]: Users´ field studies in order to analyze the appropriation of the content management system (CMS), for community oriented content production. These studies were performed by Human-Computer Interaction (HCI) experts —authors of this paper— with the participation of relevant stakeholders, who could be an iDTV application developer, an enabler of PLC return channel interactivity or an infrastructure installer [4]. Stakeholders were composed of researchers and developers from 8 countries (half European and half Brazilian) involved in the TV services deployment plus two local organizations in contact with the target users. Target users were Brazilian citizens residing in the town of Barreirinhas, a small municipality (47.728 inhabitants [5]), with a high illiteracy rate (40%). Target users belong to two different categories: primary users, meaning end users of the content, and secondary users corresponding to community content-creators.

The results of the analysis about the appropriation of secondary users for the content production will be briefly presented. They provided authors of this paper with orientation and insights of how to take into account both local content creation aspects (as the cultural aspects, the users´ expectation and interests) and the new scenario of the developing countries —as Brazil— in the phase of exploration of new ideas and solutions for future projects.

Since there are still no interactive services for iDTV in Latin America, this work can help those that can not conduct field researches to obtain such insights.

2. CONTEXT OF THE PROJECT

2.1 Project and City Background

The case study presented in this paper refers to an international project [2] funded by European Commission (EC) under the framework of collaboration between Europe and Latin America in the topic of iDTV. The main motivation of the project stands in the fact that iDTV technology presents an ideal way for bridging the Digital Divide, especially in developing countries such as Brazil where low income population does not have access to computer technology and where TV has a great penetration among the population. This project pursues the creation of a framework for enabling local communities to produce content and broadcast it through a TV channel. Citizens will be empowered with a way to participate in the process of creating and accessing digital content. For satisfying this requirement of participation, the system uses the Power Line Communications (PLC)-based networks as return channel of Set-Top Box (STB).

In this paper we will focus just on secondary users. In this town three different institutions were identified as potential Secondary Users. They were:

- SEBRAE – That is national-wide agency supporting development for small-medium enterprises and community-oriented services in Brazil with a representation in Barreirinhas
- Secretary of Education – That is a government dependency responsible of education services
- TV-MIRANTE – That is a local TV broadcaster.

2.2 Users' Field Studies and iDTV Content Creation Process

In total, there were five meetings between stakeholders and users. In the first meeting, in Barreirinhas, this research took 5 days to apply 150 questionnaires, from which 117 were received back, having an acceptable total loss of 33 questionnaires (22%). Studies had to involve an ethnographic investigation in order to observe users' experience with technology (i.e. computers, cell-phones and TV) in their daily environment. We went to users' houses and workplaces, took pictures, recorded the way they use the TV and the remote control, etc. Some interviews took place in these moments. Main aspects obtained from participants during users' field studies [6] were the following:

- Great familiarity of participants in interacting through SMS;
- Most of the manifested behaviors of the participants were based on a community-oriented lifestyle, instead of an individualistic one [7]. As for instance, Brazilian people are very motivated by social networks (see orkut success!) in order to have fun. In addition, TV is social, people see TV in group and;
- Since the people stated watching a lot of television, most of the Brazilian participants (66%) declared watching TV between 2 to 5 h/day and they preferred news (71%) and movies (62%).

In the second meeting we validated CMS requirements with 8 target secondary users. They referred to the definition of different ways to deliver contents to primary users related to its format, language, and grouping [8]. Examples of this are the existence of web templates, availability of multilingual content, visual themes, etc. We also reinforced the importance of commitment of users in the project and invited them to keep participating on this project until the end.

In the third meeting we presented a first prototype of the system for the same users. The CMS in SAMBA project is part of the Platform Domain of the overall iDTV system. A detailed description of the platform domain is presented in [3]. In particular, the CMS is responsible for the creation and management of the content and it is to provide the Web-database in which user data and content metadata are generated and stored. The CMS allows availability of information either through the Digital TV broadcasting channel or through a generic client to access to it through a simple HTML connection. To allow this, some preliminary operations are required in order to properly present these contents in a format suitable for the TV screen. In particular, the CMS communicates with the Playout Centre (the transmissor) and provides it with a specific package consisting of the broadcasted application and the associated content [1].

In the fourth meeting, a CMS training process was performed by the team (one HCI expert and one application developer). It was given to 5 secondary users about the CMS functionalities and the content creation process. The secondary user could be any type of service provider, the merchant himself, entrepreneurial support associations, such as the Small Business Support Service (SEBRAE), government agencies, etc. The Community content production process was performed only by secondary users. Stakeholders made calls in order to provide users with the technical support and to motivate them to create relevant contents.

In the fifth meeting, there was the CMS system validation. In a realistic scenario involving the mentioned local organizations, studies about the appropriation of the technology by recruited secondary users was performed. During three days, team at least of three HCI experts (designer, evaluator and psychologist) and one technician observed the execution of the activities (Creation of content, reuse of content and visualization of created content) and applied one questionnaire.

3. RESULTS

Five (5) recruited secondary users were assigned the task of creating specific content that was relevant for them. They selected the contents to be created. Then using the combination of CMS functionalities, the templates allowed the creation of iDTV pages and the association among them by forming a navigational structure. The created pages were the following: Nutrition (cashew, fruit of the cashew), Free educational courses, Fashion, Tourism, Crafts, Noise pollution, Pro-Literacy in Barreirinhas, and Information about specific local tours. The CMS had an option to reutilize other existing content available in the web through RSS feeds. However there were just a few pages updated by the three mentioned institutions, and they did not have RSS feeds. It can demonstrate TV is still the only communication midia in local communities when information needs to be announced.

According this experiment results, the CMS could not work appropriately in a local server. Secondary users used it in a remote way to create content. Such decision was related to the intrinsic limitations of the town related to low internet access. Despite this difficult, SAMBA CMS system was perceived by users as a very relevant tool for supporting the local community.

A detailed description of the users' field studies, usability tests and validation process applied can be found in papers [9] and EU-Reports [10].

4. INSIGHTS FOR FUTURE WORKS

Our experience in SAMBA has shown us how distant the dream of interactive TV is yet in Brazil. Most of the difficulties faced by the users are related to the digital gap. Even though in SAMBA, a stalemate is created. iDTV is thought as an alternative to reduce the digital gap but the implementation of that still is a challenge exactly because of the lack of infrastructure for a return channel (to provide Internet access).

Given this context, we start investigating new ways to induce iDTV in the country. One of these ways is to use conjointly the iDTV and the cell phone. From 2006 to 2009, there was a 70% raise on the amount of people who own a cell phone [11]. GPRS and 3G networks (having Internet access using a cell phone network), are provided by many private operators with a wide coverage area for this technology in the whole country. This reinforces our feeling that solutions should take advantages of the cell phone Brazilian market expansion and should allow users to make transition between media (as from TV to Cell-phone).

As an example of how this can be implemented, we have developed an application prototype for tackling a social drama in Brazil: Public Safety. Such an application is inserted into the context of collaborative mapping what has particularly been dubbed Collaborative *M*ap *App*lications (for short, WikiMapps – www.wikimapps.com) [12]. In these systems, the digital map works as a blackboard for accommodating stories told by people about events they want to share with others typically participating in their social networks. A prominent exemplar of WikiMapps is WikiCrimes (www.wikicrimes.org), a collaborative map of report of crimes. The prototype we are developing shall be used in WikiCrimes. Figure 1 illustrates a scenario, in which primary users can get complementary content, via the cell-phone, related to an iDTV talk show about Law Enforcement.

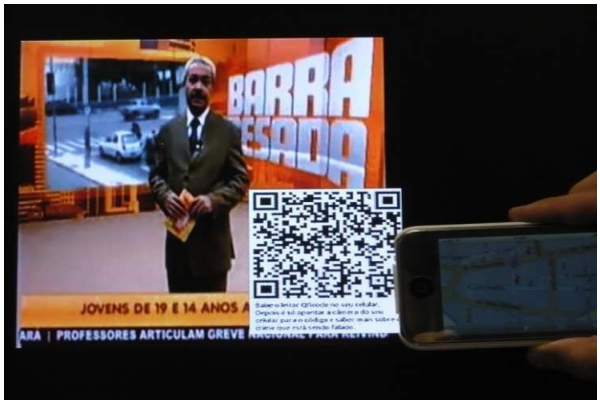


Figure 1. QR Code image seen on TV screen during a talk show about law Enforcement problems.

During the presentation of a certain news, users can require a QRCode application, installed in their cell-phone, to read the QRCode image that appears on TV screen. Via an internet connection, information about the area where the crime being reported is obtained and showed to users in their cell-phone screen.

The generation of the QR Code is done directly in WikiCrimes. Figure 2 depicts how this process is displayed in the site. The user chooses a hot spot (the shaded polygon) and generates a QR Code representing a URL that will lead to a page in which the image of the polygon is displayed.

Figure 2. QR Code image generated in WikiCrimes site.

5. RELATED WORKS

We looked for projects that address content production requirements and challenges for developing countries. Beacon [13] is a cross-cultural project on Digital TV carried out also with the support of the EC. It develops innovative t-learning services in Sao Paulo city in Brazil. In [14] there is a description of the inappropriate infra-structure of target town and solutions to introduce the technology in this context. Citizenmedia is also a multi-national project on Digital TV [15]. In all these works users' field studies were performed. Any of them discussed the results of a content production process in which can capture the richness of the technology appropriation under future perspective: sustainable project, people cultural aspects and country's economic growth.

6. CONCLUSION

In this paper we had presented the different moments of experiences between a cross-cultural development team and end users of a technological project. The main finding and contribution of this work was to identify how secondary users act in order to produce iDTV contents. Their behavior change in function of the advantages and challenges: They are satisfied when they realize relevant content for the local community can be created by themselves; They are not so satisfied when they face with technical problems, and when the iDTV style requires a new behavior of people (be more active, participative). Many small Brazilian cities are not prepared to the appropriation of this new technology. Another interesting aspect to mention is the main motivation spreads out in the country that digital inclusion will be solved via iDTV. It can be irrelevant when considering the increasing rates with internet access, cell-phones and computers for the next years. Economic and cultural aspects mentioned here can

provide useful input into the planning of design and deployment of related system, which will be developed and/or used by the same or similar sets of users.

7. ACKNOWLEDGMENTS

This work was funded by the IST EC-project SAMBA and Unifor.

8. REFERENCES

[1] Mayora, O., Fuhrhop, C., Furtado E., From Coach Potatoes to TV Prosumers: Community-Oriented Content Creation for IDTV. UCMedia 2009.

[2] SAMBA. System for Advanced interactive digital television and Mobile services in BrAzil Project. Available at: http://www.ist-samba.eu/. 2007.

[3] Mayora, O. ; Gabos D. ; Furtado, E. ; Cavaliere R.; Pascalicchio, A. ; R. Filev. A framework for community-oriented interactive digital television. infocommunications journal, v. LXIV, p. 14-23 (2009).

[4] Furtado E., Mayora,O., Fuhrhop,C., Assessments and Consensus on Users' Needs of Cross-Cultural Stakeholders: Expectations on 'Being Part of a Digital TV Project!'. To appear Euroitv'2010.

[5] Health ministery,http:/dtr2002.saude.gov.br/caadab/indicadores/maranhao/barreirinhas.pdf

[6] Vasconcelos, P. ; Fava, F.; KAMPF, T. ; Schilling, A.; Furtado, E. Ethnographic Investigational Methodology and Evaluation on Local Television Channel Creation that Allows Interaction with the Community. In: EUROiTV, 2007.

[7] Furtado, E.; Mayora, O.; Anastassova M., Kampf, T. ; Vasconcelos, P. A Investigation of iDTV User Needs in Brazilian and Italian Communities. In: workshop on challenges, perspectives and opportunities for IHC, 2007, CLIHC 2007.

[8] Furtado E., Schilling A., Fava F., Camargo L. "Promoting Communication and Participation through enactments of interaction design solutions". In ICEIS´2008, Barcelona, Spain.2008.

[9] Furtado, E. ; Kampf, T. ; Piccolo, L. ; Baranauskas, M. C. C. . Prospecting the Appropriation of Digital TV in a Brazilian Project. ACM Computers in Entertainment : CIE, v. 7, p. 10-32, 2009.

[10] Furtado, E. Mayora, O., Report about the Usability and acceptance Tests of SAMBA Applications. 2009.

[11] Holanda G., Ávila, I, Martins R., Mapping users´perspectives and outlining social impacts from digitalization of terrestrial TV in Brazil. Telematics and Informatics. 2006.

[12] Furtado, Vasco ; AYRES, Leonardo ; de Oliveira, Marcos ; Vasconcelos, Eurico ; Caminha, Carlos ; D Orleans, Johnatas ; BELCHIOR, Mairon . Collective Intelligence in Law Enforcement The WikiCrimes System. Information Sciences, p. 10-35, 2009.

[13] Beacon Project. ftp://ftp.cordis.europa.eu/pub/ist/docs/ka4/au_fp6_beacon_en.pdf

[14] Brewer, E. Demmer,M., Du,B., Ho,M., Kam,M., Nedevschi,S. The Case for Technology in Developing Regions. IEEE. Computer Society. 0018-9162/05/. Pg 25-38 June, 2005

[15] Citizen Media, http://www.ist-citizenmedia.org

The Digital TV as another solution to educate in isolated areas in the Amazon State, Brazil

Lílian Simão Oliveira
ICMC - Universidade de São Paulo
Trabalhador são-carlense Avenue, 400 – Downtown
São Carlos – SP, Brazil
Postal Code: 668 - CEP: 13560-970
Phone: 55 (16) 3373-9700
lilianso@icmc.usp.br

Maria da Graça Pimentel
ICMC - Universidade de São Paulo
Trabalhador são-carlense Avenue, 400 – Downtown
São Carlos – SP, Brazil
Postal Code: 668 - CEP: 13560-970
Phone: 55 (16) 3373-9700
mgp@usp.br

José Pinheiro Queiroz-Neto
IFAM – Instituto Federal Amazonas
Av. Gov. Danilo Areosa, SL01
Distrito Industrial
Manaus – AM, Brazil
CEP: 69075-351
Phone: 55 (92) 36133530
pinheiro@ifam.edu.br

ABSTRACT
Several studies show that the distance education is a viable way to offer education to places that have difficulties in access to quality education. In the Amazon State, Brazil, some projects have gained notoriety for their emphasis on providing effective education at various high school and higher education levels in remote cities. We aim at providing mainstream inclusion to people who live in hard-to-reach cities. The aim of this project described in this paper is to investigate alternatives to use an e-learning management system within an interactive digital TV scenario. We plan to use the TIDIA-AE platform, which allows synchronous and asynchronous collaborative distance learning supported, integrated with the interactive digital TV platform under development in Brazil. As a result, we are providing a proof-of-concept prototype tool to be used by institutions in their educational activities. One of the tasks is the customization of the TIDIA-AE software platform taking into account some requirements for its use in a region as the Amazon, which presents very limited good quality network access. Our proposal takes advantage from the quality of image that TV provides and investigates the advantages of interactive digital TV. The work includes: investigation of the requirements, proposal and evaluation of alternatives, development of the proof-of-concept prototype and experimentation within potential scenarios. The work will be summarized in a report discussing the feasibility of using TIDIA-AE and Interactive Digital TV as integrated platforms to support the distance education in the Amazon state. The project is work in progress, and we plan to submit it to the pilot classroom group for tests in mid-2010.

Categories and Subject Descriptors
D.3.3 [**Programming Languages**]: Language Constructs and Features – *abstract data types, polymorphism, control structures.* This is just an example, please use the correct category and subject descriptors for your submission. The ACM Computing Classification Scheme: http://www.acm.org/class/1998/

General Terms
Your general terms must be any of the following 16 designated terms: Algorithms, Management, Measurement, Documentation, Performance, Design, Economics, Reliability, Experimentation, Security, Human Factors, Standardization, Languages, Theory, Legal Aspects, and Verification.

Keywords
Distance Learning, Interactive Digital TV.

1. INTRODUCTION
There are numerous difficulties in achieving quality education in the Amazon State, Brazil. Correa [4] listed difficulties such as lack of sanitation, treated water, health facilities, and electricity (in some remote regions, electricity can only be accessed by batteries or oil generators). He also observed that access to school in this region may be only up to the 4th grade school level.

Since 2004, innumerous projects have been developed in order to expand the education available in the countryside, including distance-learning based projects. One important example is the project introduced by the State Secretary of Education (SEDUC), which has recently introduced distance learning for high school using new technologies. The project, employing digital TV technology via IP, allows a thorough integration of a TV studio infrastructure and live online classes which include participation via chat, e-mails as well as videoconferencing. Platforms such as the one deployed by SEDUC are expensive and hard to customize. For instance, accessing students' data with respect to their interaction with the system is not provided and hard to implement.

Open source state-of-the-art distance-learning based projects based on traditional Internet are difficult to be deployed given the low reach of Internet in the State as a whole. As a result, low quality interaction is achieved by having some type of broadcast (e.g. Terrestrial TV, very common in Brazil) used in combination with feedback via telephone or even fax.

Investigating the current scenario deployed in the Amazon State, we observed the problem of the low level interaction allowed by the available infrastructure. For instance, there is a necessity for alternative ways to expedite the handling of all submitted questions from the participants, so that all questions can be answered in a timely fashion. It is important to highlight that, as observed by Landim and Ribeiro [9], the delay in providing feedback to students in distance learning environments influence student motivation and drop out.

In the remaining of this paper, we detail our work in progress in which we investigate the opportunity of integrating a state-of-the-art open source Internet-based infrastructure, the TIDIA-AE (*Tecnologia da Informação no Desenvolvimento da Internet Avançada – Aprendizado Eletrônico* – Information Technology in the Advanced Internet Development- e-Learning) platform [17], with the novel terrestrial-based interactive digital TV (iDTV) adopted in Brazil, based on the Ginga-NCL middleware [11].

2. RELATED WORKS

The literature advocates the use of learning objects in the context of interactive TV in a way that combines education and entertainment [10]. Some iDTV distance education applications, mainly interactive games and quizzes are described by dos Santos et al. [7]. An outstanding argument in this work is the provision of educational interactive documents as value added complement to the TV program; similar observations are pointed out by Soares and Souza Filho [16]. The research reported by Im et al. [8] presents a system for transparent conversion of web-based learning objects into iDTV applications, but without considering divergences in usability between Web and DTV. A feasible way to fight these disadvantages is presented by Lopez et al. [10], which includes the content authoring process, implying better control on the format of the learning objects. Common to all those related works is the focus on the MHP middleware, employed in the European DTV system, which makes them not immediately portable to the International System for Digital Television (ISDTV) that is the Brazilian Digital TV standard to terrestrial mode.

Cesar [3] comments the TV as social role and said the importance to develop thinking in the user that will consume that content. The research of [2] [12] shows that in Brazil the same path to transform learning objects from web platforms for digital TV is the way to the automatic generation of content for TV, and [12] test a prototype but using only local interactive. This project propose using the return channel but with restriction like a dial-up internet.

These presented papers shows that DTV for learning purposes is an area expanding and demand surveys. Transforming educational content from platforms LMS (Learning Management System) is a viable alternative to generation of DTV learning objects. This requires learn to use the new media, considering interaction and interface adaptations. Our project proposes to use the DTV as another tool interaction between students and teachers in distance education, using the Brazilian standards middleware – Ginga. Developing user-centered design and concerned with the usability of the system. With aim to provide tools to increase the education in isolated areas like Amazon state.

3. THE PROJECT

The main objective of the project is to investigate and implement a distance learning solutions applicable to the Amazon State using interactive digital TV.

An open platform was used as a base for in class education and distance learning, the TIDIA-AE [6,7]. We are designing and experiencing new modules so that they can allow the integrated use of the digital TV with an open source Internet-based platform: some modules are currently under development by other research groups University of São Paulo, Brazil.

The Digital TV module will be designed with more emphasis as a tool for the kind of teaching we envision for the State. The fact is that many cities in the remote areas of the state face difficulties in terms of Internet connectivity, which does not allow the use of Internet-based audio and video communication. However, we understand that the quality of educational process depends on continuous and bidirectional communication and, without Internet connectivity, the distance-learning approach reverts to the old model of independent study via mail in which the student becomes autonomous and isolated, and tends to drops out.

Our work is divided in six stages as follows:

1. Survey data about distance learning in the state
2. Elect one distance-learning ongoing effort as case study
3. Investigate the use of the TIDIA platform environment elected for the case study
4. Investigate alternatives of integrating the SBTVD and TIDIA platforms (requirements engineering)
5. Design and implement of a proof-of-concept prototype, in cooperation with the TIDIA and SBTVD teams
6. Experimentation with potential users.

3.1 Survey data about distance learning

We have concluded Stage 1: the main projects that have been surveyed:

A) Undergraduate and graduate courses offered in the remote areas were developed by the University of the State of the Amazon (UEA). Conventional TV was used for broadcast and students' questions were answered via phone, fax and e-mail. Started in 2002, it is ongoing. It has several courses and the first attend every city in the Amazon state.

B) High school distance learning, offered by SEDUC using IPTV technology. The hardware and software infrastructure allows efficient tools for communication, control and follow-up. This program was implemented in 2007 and the first group will graduate late 2009.

C) Vocational technical level course in Tourism and Environment are offered by the Federal Institute of the State of the Amazon (IFAM). These courses will be offered in cities such as Barreirinha, Eirunepé, Manaus, Tabatinga and Tefé, starting in the first semester of 2010. These course will offer using only Web and the Learning Management System used is the Moodle.

After this survey we observe that the TV has important hole in distance education in Amazon. With this technology is possible attend every cities in the state. It has too the IPTV project that provide the class with great image and audio quality. Using this technology it can be possible the students interact in real time using chat and videoconference to ask the teacher who stay in studio in Manaus – Amazon capital.

It was also observed that the Internet can be used with a teaching tool but not exclusively, because it does not has a good quality and it is unstable, less than 524kbs in some cities in state, making the connection impossible at times due to technical problems.

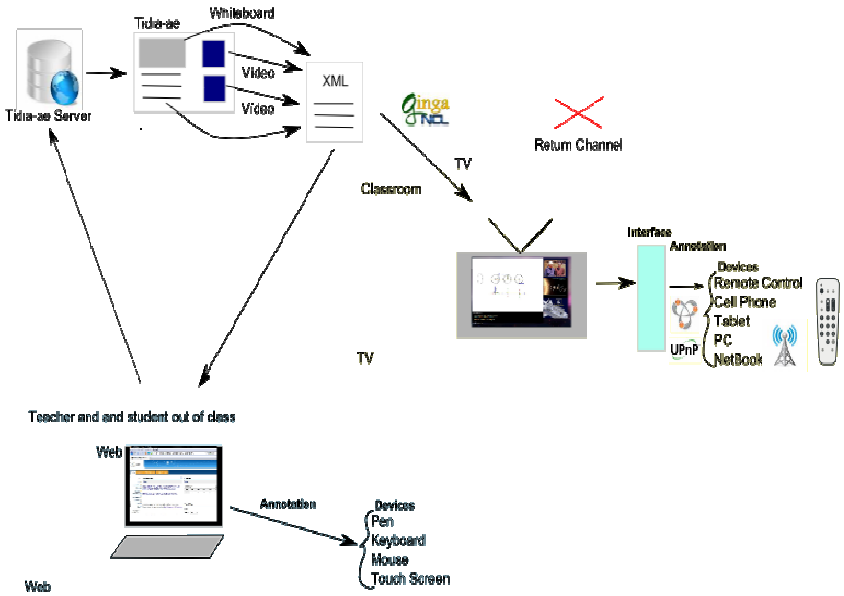
3.2 Elect one distance learning project

The IFAM Project using only Web to provide content class. After the survey was observed that the internet has serious problems in state. Because this, the IFAM can be offer its courses to all cities of Amazon due the technique problems that can be impossible the students follow the lessons offered.

In Stage 2, The IFAM Project was chosen due to its flexible characteristics and for being a project that gives the necessary inputs for the research development.

3.3 TIDIA-AE

The TIDIA-AE (*Tecnologia da Informação no Desenvolvimento da Internet Avançada – Aprendizado Eletrônico* – Information Technology in the Advanced Internet Development- e-Learning) is a LMS developed by the São Paulo Universities. It is an



Picture 1 – Scenario used in this project

ambient that has tools that support the presence and distance education. These tools can be: - synchronous, this mean that the student talk with the teacher in real time like chat; - asynchronous, that every one use the tool in your time, like forum that the student write your question and when the teacher entry in the ambient and read the question, write the answer; - collaborative, that every participant have the same permission and can write and create content independent if he is a teacher or a student like a wiki.

The Tidia-ae project starts the development in 2004 and concluded in 2009. It was initiated by FAPESP (the State of São Paulo Research Foundation) with the main objective of developing an e-learning system that can explore the potential of Advanced Internet and can provide support to different educational context needs. The project integrates various research groups and is guided by the following main goals [1] : (i) to provide a flexible system that can be assembled from a set of selected tools and functionalities in order to satisfy various learning scenarios as well as to adapt the existing tools and functionalities to the needs of the various research groups; (ii) to promote the reuse of tools and/or functionalities developed by TIDIA-AE; (iii) to promote interoperability with other electronic learning systems; (iv) to define a process of collaborative development among the different research groups.

In Stage 3, from the TIDIA-AE infrastructure, we have elected the following tool: text-based Chat, videoconference, whiteboard. We will exploit audio and video tools as authoring aids for instructors and students to produce interactive digital TV programs. The information produced by all these tools will be processed by new authoring tools so as to automatically generate multimedia documents appropriated to the interactive TV.

3.4 SCENARIO

The Picture 1 describes the scenario that we are using in this project. It has a Tidia-ae platform that storage in a server. The Tidia-ae can be different tools as a wiki, chat etc, in this project we will work with chat, video conference and whiteboard. These tools are describe in a XML archive that specifies every media used, when is used and who used it. Using this archive is possible convert the session create in Tidia-ae to Digital TV, using the Brazilian middleware – Ginga – and be present in Classrooms located in cities of the Amazon state. In this classroom every student can watch the teacher explain the content with good image quality. The interaction with the content can be done using the remote control, where who use it is the tutor that is the person who orient the class where there is no teacher.

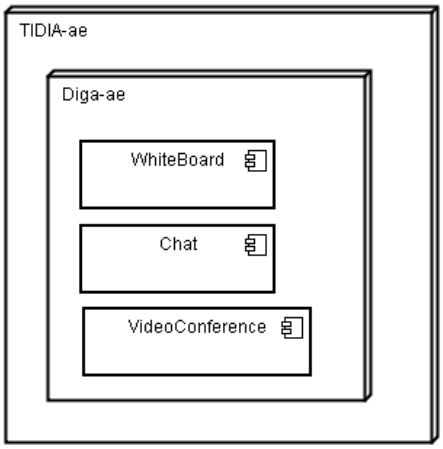
In another stage of this project, we can investigate the possibilities to use the others devices that can individualize the interaction with the TV like PC, Cell Phone, Netbook etc, using a local network with the set top box or Digital TV.

The internet in Amazon State in most places isn't good. The internet bandwidth is less than 524kps. In this cities the interactive will be only local, this mean that the questions and answers done in classroom don`t will be send to the server. To minimize this problem will be implementing and test the return channel using the dial-up internet. Than the file send in return channel to the server will be only text, the student can write his question using the chat tool.

The class can be recorded by the teacher in a studio using a video conference session. It is necessary an internet that support this service then will be done in a capital of Amazon or other cities that the internet is good. To help the teacher explain a content he can use a whiteboard. It can upload a class slides and possibility the teacher annotated the important things in a specific slide or he can use the chat too. This session with the video and audio of the video conference synchronized with the whiteboard and the chat is saved in a Tidia-ae and can be converted to the Digital TV using the Brazilian middleware to DTV – Ginga – to be presented to the students in classroom in a specific time. The document describe in XML will be converted and done the adaptation to NCL to be presented in a DTV.

Additionally the classes can be accessed by the students that have internet at home or in lan houses, just accessing the Tidia-ae website.

In Stage 4:



Picture 2: The Diga-ae tool of TIDIA-AE

The Picture 2 show that the LMS Tidia-ae can have a different tools combination and in this project we choose the Diga-ae tool that is the group of tools like whiteboard, chat and videoconference where all these tools are synchrony and collaborative.

The Picture 3 shows the information flow from the student access using your computer at home until view the class using the DTV and return data, using the return channel. The use of this technology would be the teachers scoring one day to take the doubts expressed by students. At the appointed time they would access the Diga-ae using the Internet and would record the class collaboratively using video conferencing, chat and whiteboard. This session could be subsequently accessed by students accessing the Tidia-ae platform or watching this on DTV. In this second case is necessary an adaptation and conversion of media from the web. This conversion consists in obtain media descriptions in XML documents, these documents are subsequently converted to the NCL format after passing through a module that will be entrusted with this task. This module was developed using Java technology. The NCL document is generated in accordance with the specifications of the Brazilian Digital TV middleware - Ginga, more specifically the Ginga-NCL. After this step the document may be filed in the DTV, at the class scheduled. The interaction in the classroom might be using the chat that could be accessed using the remote control or keypad. Because are many students and only one device interaction, the tutor would be the person responsible for input data. This person is responsible for intermediate students and teachers, facilitating the conversation between them. These data generated in the classroom would be sent to the server's Tidia-ae through an internet connection dial-up. On the server that the data would be consolidated and updated. When a student uses the Internet to access the system, he can view the questions generated in class televised through the chat logs.

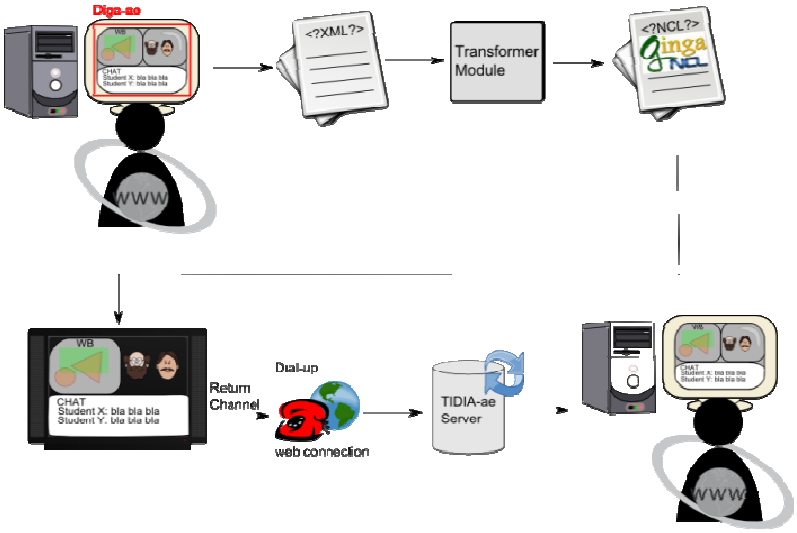
Our work on Stage 5 should implement the functions describe in this scenario and as result in proof-of-concept prototypes of those new authoring tools.

3.5 TESTS

Concerned about the acceptance and functionality of the prototype generated will be developed usability tests to validate the implementation.

Initially be implemented heuristic [13] to validate the usability of the system. These are made with professionals who can be the domain experts, i.e. distance learning or usability and interaction experts. This initial test will be developed with aim to find critical issues before applying to the potential users. Once implemented the corrections of the main problems encountered in the heuristic test will run a test with potential users, in a pilot class with students of IFAM course.

After these tests we expect know if the prototype has a good acceptance and it can be used with success in distance learning projects.



Picture 3: The information flow in proposed system

4. FUTURE WORKS AND CONCLUSION

We intend to finalize the implementation of the architecture proposed in this paper, and test with the real student in a pilot classroom. The real tests are important to evaluate the usability of the class using the DTV and known the problems create in automatic adaption of a LMS to a Digital TV.

By exploiting state-of-the-art efforts in terms of Internet-based learning management systems and interactive TV, at the end of this project we aim to have experimented with a proof-of-concept prototype tool to support distance-learning that takes into account the unique characteristics of the Amazon State in terms of needs and available infrastructure. We expect our work will have impact in several education-related problems in the State of Amazon in Brazil.

5. ACKNOWLEDGMENTS

We thank FAPEAM, FINEP, CAPES, FINEP and FAPESP.

6. REFERENCES

[1] BEDER, D. M.; CASAGRANDE, L. and RUBIRA C. *TIDIA-Ae Project Software Architecture Proposal*, Internal Document, TIDIA-Ae Project, São Paulo, Brazil, 2004

[2] CELES, C. S. F.; SOUZA, C. T. (2007). Estilos arquiteturais de software na construção de objetos de aprendizagem para a tv digital interativa. *XVIII Simposio Brasileiro de Informática na Educação (SBIE), pages 1–10*

[3] CESAR,, P.; BULTERMAN, D.C.A.; SOARES, L.F.G., (2008) - Human-Centered Television - Directions in Interactive Digital Television Research - in *ACM Transactions on Multimedia Computing, Communications and Applications*(TOMCAP), article number 24, 2008.

[4] CORRÊA, Sérgio. Trabalho, saberes, identidades e tradições nas comunidades rurais riebirinhas. In: OLIVEIRA, I.A. (Org). *Cartografias Ribeirinhas:Saberes e representações sobre práticas sociais cotidianas de alfabetizandos amazônidas*. CCSE-UEPA, 2004.

[5] CUNHA, Conceição Maria da. Discutindo Conceitos Básicos In Educação de Jovens e Adultos. *MEC – Secretaria da Educação à Distância*, 2001.

[6] DE OLIVEIRA, J.A., WAISMAN, T., SANTOS DA SILVA, F., HORBE, B., MAIA, O.B., AND MORENGHI, F. 2006. T-learning in Amazon: The interactive education project. *In Proc. EuroITV 2006*, 290-295

[7] dos SANTOS, D. T.; VALE, D. T. and MELONI, L. G. P. Digital tv and distance learning: Potentials and limitations. In *Frontiers in Education Conference, 36th Annual*, pages 1–6, March 2007. DOI = http://dx.doi.org/10.1109/FIE.2006.322670

[8] IM, S. ; LEE, S.;WU, X.; LEE, M. and HWANG, D. - Design and Implementation of SCORM Content Conversion System for DiTV. In *ICHIT '06: Proceedings of the 2006 International Conference on Hybrid Information Technology*, pages 679–684, Washington, USA, 2006.

[9] LANDIM, Cláudia e RIBEIRO, Antônia Maria Coelho. A elaboração de projetos, Unidade 5, SENAC 2000. *Série Educação à Distância*.

[10] LOPEZ, M. R.; REDONDO, R. P. Diaz, VILAS, A. F. and PAZOS ARIAS, J. J. Entercation: engaging viewers in education through TV. *Computers in Entertainment (CIE)*, 5(2):7, April 2007.

[11] *Middleware Aberto do Sistema Brasileiro de TV Digital (SBTVD)*. http://www.ginga.org.br/. Visited on 23/03/2009.

[12] MONTEIRO, B. de Souza (2009). Amadeus-tv: Portal educacional na tv digital integrado a um sistema de gestão de aprendizado. Master's thesis, Universidade Federal de Pernambuco.

[13] NIELSEN, J. and MOLICH R. - Heuristic evaluation of user interfaces. In *VIII SIGCHI conference on Human factors in computing systems: Empowering people*,1990.

[14] OLIVEIRA, L. S. ; PIMENTEL, M.G. ; QUEIROZ-NETO, J. P. . New challenges to educate in isolated areas in the Amazon. In: 39th IEEE Frontiers in Education, 2009, San Antonio. 39th IEEE Frontiers in Education. Champaign, IL : Stipes Publishing LLC, 2009. p. WC1-1-WC1-6.

[15] STIUBIENER, I. ; RUGGIERO, W. V. ; SILVEIRA, R. M. ; KOROLKOVAS, I. ; KOPP, S. ; BAPTISTA, C. M. . Evaluation of a Real Network Experiment in a Remote Environment. In: *III Workshop TIDIA*, 2006, São Paulo. p. 121-123.

[16] SOARES, L. F. G. and SAOUZA FILHO, G. L. Interactive Television in Brazil: System Software and the Digital Divide. In *Proceedings of the EuroITV 2007 Conference*, Amsterdam, Netherlands,2007.

[17] *The TIDIA-Ae Portal at Universidade de São Paulo*. http://tidia-ae.usp.br. Visited on 23/03/2009.

WORKSHOP 2

Methods for User Studies of Interactive (TV) Technologies

Editors:

Dr. Paul Marrow (BT Innovative & Design, UK)
Dr. Lydia Meesters (Eidhoven University of Technology, Netherlands)
Prof. Marianna Obrist (ICT&S Center of the University Salzburg, Austria)

SocialMedia Experience and Design Lab: Using Advantages from Different LivingLab-Approaches

Jan Hess
University of Siegen
Adolf-Reichwein-Straße 2, D-57068 Siegen
+49 271 740 2098
jan.hess@uni-siegen.de

Corinna Ogonowski
University of Siegen
Hoelderlinstr. 3, D-57068 Siegen
+49 271 740 3265
corinna.ogonowski@uni-siegen.de

ABSTRACT
Within this paper we describe our approach called SocialMedia Experience and Design Lab (SMEDL). Within SMEDL we operationalize different levels of user involvement in real-world settings (at home and on the web) as well as in an artificial lab environment. Because the different approaches have their strengths and weaknesses, we highlight the meaning of a method-mix around a Living Lab as our core concept. We also describe methodological issues that have to be considered in further work.

Categories and Subject Descriptors: H.5.1 [Multimedia Information Systems]: Methodology

General Terms: Design, Human Factors, Theory

1. INTRODUCTION
Social Media has become an actual phenomenon in our modern world. One of the most commonly known definitions of Social Media in its sense of Web 2.0 is based on Kaplan and Haenlein: 'Social Media is a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content' [6]. Even though the definition is generally accepted, it has to be seen from a broader sense of media use. Based on this assumption we started a large-scale research project called 'SocialMedia' at the University of Siegen, which aims to develop new cross-platform community concepts for domestic environments based on TV-, PC- and Mobile applications. In the project we develop several use cases related to SocialTV research (e.g. chat, annotation and social recommendations). By developing new applications in a Living Lab context, we will support communication and shared content processes in existing communities, as well as provide support to community building.

The participation of users in the design and evaluation process of new technical artefacts is a fundamental approach in building usable and acceptable applications and services. We chose a Living Lab approach in order to bring users, businesses and technology together into a development process that establishes real life environments [7]. According to Eriksson et al [1] such an approach supports long-term cooperation, co-creative research and the continuous involvement of users from an early stage. Living Lab approaches in Home Entertainment can be divided into two different lab structures. One structure uses as an artificial environment, in which real life home structures are simulated in test centres, such as the PlaceLab of the MIT [5]. Such a lab structure offers a controlled environment with multi-observation possibilities over a long period of time, which supplies a wide array of quantitative data. The other Living Lab structure uses real-life households as testbeds without creating an artificial setting. By utilizing user participation in innovative development processes, the project 'iiTV@Home: Field trial in Salzburg' uses such a lab concept in order to gain a deeper understanding of social dimensions in Home-IT [8].

Table 1. Pro and cons of the two Living Lab approaches

Lab approach	Praxis approach
+ quite structured procedure with a formalized output + large number of test persons + multi-observation (several video sources, observation) ± predominantly quantitative data (questionnaire, logging) - artificial lab structure - short-term evaluation	+ non-artificial lab structure + understanding praxis & context + long-term evaluation ± predominantly qualitative data (interview, diary, probes) - multi-observation difficult - selection process difficult - time- and work-consuming

As described in Table 1, both approaches have their strengths and weaknesses. Based on the experiences of the two lab structures, we decided to use a method-mix for our SocialMedia project to take advantage of both. The lab is called SocialMedia Experience and Design Lab (SMEDL), because user experience and design will go hand in hand when creating new applications and services for Home-IT. With this approach we try to find answers to the following questions: How do we get users involved in the design and evaluation process? Which methods are strong enough to help us ascertain a deeper understanding of media use? And how can we combine these methods in order to approach the challenge, which is affiliated by the Home-IT context?

2. OUR APPROACH
Within the SocialMedia project, SMEDL plays the role of an infrastructural toolkit that collects new ideas from the users (Open Innovation) and evaluates early prototypes in the domestic environment. This infrastructure will enable us to conduct qualitative as well as quantitative methods. Because the aim of SocialMedia is the development of new community tools, SMEDL will offer a test bed with different clusters of households. Participants who know each other will form a cluster to evaluate concepts for existing communities. At the same time, SMEDL includes separate clusters of households with the aims of evaluating concepts for community building and of getting new contacts. Thereby the lab will consider households with different levels of technical expertise (Lead User as well as persons with less domain knowledge). However, the evaluation of new prototypical concepts is not the only goal of our research. We also want to use our lab to explore media usage and its possible

changes over a long period. In order to operationalize the several issues, we structured SMEDL in the following way: **SMEDL.Local** is the core of our Living Lab approach. For this part we will choose ca. 20 households in the city of Siegen and the surrounding area and let them continuously participate. These households will get new hardware and software to test and evaluate early prototypes in practical settings over a longer period of time. **SMEDL.Stat** is a stationary laboratory room at the University of Siegen, which is built exactly like a standard living room. Within this environment we can measure user feedback and quantitative data on a very exact level. **SMEDL.Global** is an environment, which supports the innovation process by existing online communities (e.g. Media Center communities). On these platforms we will evaluate our concepts in a distributed manner.

While each part of SMEDL is an interesting lab approach on its own, the full value of our concept will be reached by running the three parts in parallel. By continuously involving and observing users in real world settings (SMEDL.Local), measuring their behaviour in a standardized environment (SMEDL.Stat), and by gathering feedback from a broad online community (SMEDL. Global), we expect to be able to evaluate new prototypical concepts in a full scope of richness.

3. CURRENT STATE

The work around SMEDL is embedded in a long-term research strategy. Our agenda on the one hand includes research on prototypes for supporting communities in domestic environments (e.g. [2]). On the other hand we focus on methods and concepts to explore social practise in the field (e.g. [4]) as well as involve online user communities in the design of new products [3]. The experience gathered in the past helps us structure and plan for SMEDL in a realistic way. As one of the first activities, we started to implement SMEDL.Stat and the build up process for SMEDL. Local. While SMEDL.Stat is already running, the setup of SMEDL.Local is an ongoing process. Right now we are performing ethnographical work in the field to explore the current state of media usage and to identify existing networks of participating households. When the build up process on SMEDL.Local is finished, we will start to realize the community portal for SMEDL.Global. Because a critical mass of users within SMEDL. Global is necessary, we want to access existing community portals, which already are established.

4. DISCUSSION

Choosing the right method for a given research issue is not a trivial task. Should we evaluate a concept in the artificial lab, in praxis or in both? Should we explain some details of the concept at first or should the user explore it with no additional explanations? Is it more valuable to conduct interviews with single persons or with all members of the household at the same time? Can we gather more insights with rather classical ethnographic methods (e.g. observation) or should users document on their own (e.g. by using diaries or probes)? From our previous experience we can state that the right choice is a trade-off between several conceptual (new ideas vs. evaluate concepts), technological (easy to use vs. complex technological preconditions) and organizational (e.g. budget and manpower) aspects.

Depending on the concrete design goal, the best choice of methods differs from case to case. A potential source to find new ideas is the Lead User approach [9]. Such users are very experienced in a domain, use an artifact in a regular manner and therefore have good ideas for improvements. We already gathered positive results with community participation based on a virtual platform [3]. In this study interested users could contribute in a distributed manner to define the functionalities of a new media center system. Even if a lot of different ideas and improvements were brought in this way, the process of evaluating new concepts becomes difficult. The procedure will profit from personal contact between user and designer. The common ways to evaluate concepts on site are controlled lab studies and workshops with the users. By conducting interviews, observation, logging etc., numerous of fundamental insight can be gathered. However, even such an approach is limited in different ways (see also Table 1). Many of details will remain unanswered: Is the concept accepted in praxis? Will the user make value of it over longer periods of time? In which way and in which context will it be used? How does the concept change the practise of the user in everyday life? Such insights will observable only in real-life settings. Because all of the introduced lab structures – virtual lab, artificial lab and praxis – have their value for a broad understanding of design issues, we operationalize all three of them for our SocialMedia project as described in the previous sections.

5. REFERENCES

[1] Eriksson, M., Niitamo, V.-P. and Kulkki, S. 2005. State-of-the-art in utilizing Living Labs approach to user- centric ICT innovation - a European approach. Working Paper.

[2] Hess J. 2008. Community Awareness in SocialTV Environments. Second Workshop Social Interactive Television. EuroITV (Salzburg, Austria, Jul. 3-4 July, 2008).

[3] Hess, J.; Offenberg, S. and Pipek, V. (2008): Community-Driven Development as participation? - Involving User Communities in a Software Design Process. In: Participatory Design Conference (Bloomington, Indiana, USA, Oct. 1-4, 2008).

[4] Hess, J. and Wulf, V. (2009): Explore Social Behaviour around Rich-Media: A Structured Diary Study, European Conference on Interactive TV (EuroITV 2009).

[5] Intille, S.S., Larson, K., Beaudin, J.S., Nawyn, J., Mungui Tapia, E. and Kaushik, P. 2005. A Living Laboratory for the Design and Evaluation of Ubiquitous Computing Technologies. Proceedings of the Computer Human Interaction (Portland, Oregon, USA, April 2-7, 2005), 1941-1944.

[6] Kaplan, A. M. and Haenlein, M. (2010): Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 53 (2010), 59-68.

[7] Niitamo, V., Kulkki, S., Eriksson, M. and Hribernik, K.A. 2006. State-of-the-Art and Good Practice in the Field of Living Labs. In Proceedings of the International Conference on Concurrent Enterprising (Milan, Italy, Jun. 26-28, 2006), 349-357.

[8] Obrist, M., Bernhaupt, R. and Tscheligi, M. 2006. Interactive Television for the Home: An ethnographic study on users' requirements and experiences. In Proceedings of the EuroITV (Athens, Greece, May 25-26, 2006), 349-358.

[9] von Hippel, E. 1986. Lead Users: A Source of novel product concepts. Management Science 32 (1986), 791-805.

Using a Method Mix to Overcome Limitations in Studies of Technology Usage in the Home

Michael Pirker
IHCS-IRIT
Université Paul Sabatier Toulouse
118 Route de Narbonne
31062 Toulouse, Cedex 9, France
Michael.Pirker@irit.fr

Mahmoud Abdellatif
IHCS-IRIT
Université Paul Sabatier Toulouse
118 Route de Narbonne
31062 Toulouse, Cedex 9, France
Mahmoud.Abdellatif@irit.fr

Regina Bernhaupt
ruwido user experience research
Köstendorferstr.8
5202 Neumarkt, Austria
Regina.Bernhaupt@ruwido.com

ABSTRACT
A variety of methods has already been presented to investigate the user's technology usage around interactive television and to capture the user's experiences in different usage contexts. The selection of the appropriate method(s) to be used is dependent on various factors like research goals and perspectives, hypotheses available, decisions whether qualitative or quantitative data should be gained, understanding of methods and the insights that thus can be gained, and a clear research framework. Additionally, the application of the method is influenced by factors like the context in which the evaluation takes place and the application domain. In this position paper we argue that the limitations set by various methods can be overcome by using a method-mix and a triangulation of methods.

Categories and Subject Descriptors
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Measurement, Experimentation, Human Factors, Theory

Keywords
Home study, qualitative method, quantitative method, field study, lab study, limitation, method-mix.

1. INTRODUCTION
Studying technology usage in the home has become an emergent field in the area of human-computer interaction. The ongoing change from analog to digital TV and the continuous growth of IPTV offers almost all over the world were two of the major driving forces for the increase in technology usage studies. A wide variety of methods and methodological approaches was used to understand what users might want to use, what they are interested to use and – from a commercial viewpoint – what technologies and services they would be willing to pay for. Investigations on technology usage today reach far beyond a simple understanding of who is using what and when, but do focus on a variety of software quality aspects like usability, security, privacy, trust and, more recently, on the general user experience.

What we have learned in the last 20 years of studies in the home was the general need to adopt (social science) methods to help investigate the home context. First methodological adoptions have been presented in the late 1990ies, using ethnographical inspired methodologies to understand users' technology usage.

About 10 years later, the usage of ethnographic methods and adoptions has become standard to investigate media usage and to gain insights into everyday user behavior. To fully understand current technology usage and the possible take up of new forms of IPTV (Internet Protocol Television) offers, researchers used field studies (using prototypes or real systems) combined with observations and interviews; furthermore, a variety of usability evaluation methods was used including empirical or experimental methods, observational methods, query methods and even the use of physiological monitoring.

From a researchers perspective we currently have to face the following problems: (1) Using (ethnographic) methods that have been adopted to investigate the home context, how can we be sure that we are able to understand, investigate and gather all aspects of technology usage? (2) How can we understand the limitation of these methods and (3) is there a way to overcome these limitations?

In order to investigate these three questions, we will present a brief overview on methods currently used to investigate interactive TV and IPTV within the home context in the following. We then present the problem description in more detail and propose a methodological solution for the problem described.

2. STATE OF THE ART
Understanding technology usage and especially investigating usability and user experience as important factors of technology usage has been performed using a variety of different approaches, ranging from focus group sessions, collaborative and participative design, online surveys and questionnaires to the logging of data, usability evaluations and field trials.

One way to gather qualitative data about users is to use an ethnographic approach, in which researchers must become "immersed" in the everyday world of their research subjects in order to reach a "qualitative understanding" of what happens in the social setting [7]. This approach can be enhanced by several methods to gain additional insights and material, like video ethnography [7] or the cultural probing approach with different variations of it like the creative or playful probing approach to support the self-observation of the participants with some pre-structured material and include also children [2].

Rice and Alm [6] used two kinds of usability methodologies for their design of innovative user interfaces for elderly television viewers: interactive theatre and wizard of oz testing. Harboe et al. evaluated their various devices for social television in field trials [5], and also focus groups [4]. For long term media changes in-depth interviews are used (e.g. [1]).

3. PROBLEM DESCRIPTION

The variety of methods used to investigate technology usage shows how difficult it is to select the right kind of question. When choosing a method we have to take into account the following aspects:

First, the *research perspective*: We have to differentiate between: (1) research that focuses on exploring a new area, with the goal to build a theory based on the gathered data, and (2) research that tries to answer a given research question (hypothesis) with an already well formulated theory.

Second, the *research question*: Depending on the research questions and research goals we have to choose a different method or more likely a set of methods. A possible strategy to improve the validity of the methods is to use a set of methods to answer the same research question (thus different methods should have the same results, and therefore allow improving the validity of the data). Especially for applied research it is important that the actual research problem defines the choice of the method. What we have seen in our practical work with industry is that sometimes the knowledge on how to conduct a method is the selection criteria.

4. APPROACH CHOSEN

To understand the impact of the method choice, the current limitations of methods and how to overcome these limitations, we have started to set up a method classification system on Usability and User Experience Evaluation in Non-Traditional Environments (USENTE) [3]. USENTE should allow us to summarize the different contextual factors that can influence the application of a method.

The USENTE Framework proposed to classify the choice of a method taking into account research perspective and research question. The research question could be typically categorized by:

(1) Domain defined by either the physical context (e.g. home or mobile) or defined by special attributes of the system (e.g. multimodal interaction for safety-critical applications).

(2) Usage Aspects (typically user centered factors like usability and user experience, but today also in combination with other factors like trust, security, privacy)

(3) Influences on the user (if the technology is for example tiring the user or if the software is intended to entertain the user on a long term)

(4) Variety of user groups.

While this framework is not subject of this position paper, we want to demonstrate how such a classification can help to select a good set of methods for a research problem.

Keeping in mind the framework, and focusing on our research goal, which is to investigate and understand which new forms of multimodal interaction in the living room would be helpful to support new IPTV based entertainment services, the following factors have to be addressed:

(1) Research perspective: is there a theory about multimodal interaction or are we interested in understanding current usages to build a theory on these insights and observations?

(2) Our research question is to gain insights on how privacy, security and trust is influencing the usage of multimodal input devices – what factors do we have to take into account when choosing our set of methods?

Selecting the evaluation method, we have to take into account:

(1) The domain: In the home domain it is difficult to have researchers participating in everyday life. Evaluation methods will be based on self-observation or observation by technical equipment (cameras).

(2) Usage Aspects: we are interested in a multitude of factors, thus a method mix is necessary to address all these factors. The self-observation approach should be complemented by interviews and questionnaires.

(3) Influences of the user: to understand possible different contextual influences on the user, it would be helpful to investigate these aspects in user studies in the field.

(4) Variety of user groups: to reach a broad variety of user groups, the self-observation method must be performed in a broad variety of households, with at least 10 members representing each evaluated user group.

5. SUMMARY

To sufficiently research the topic of iTV technology, including emerging mobile services and different context and target groups, a method mix of quantitative as well as qualitative methods seems appropriate, taking into account and appropriately addressing both the research perspective and research questions. Classifications, as briefly shown with our framework, can help to select a good set of methods to properly address a research problem. Additionally, combinations of laboratory and field studies could provide (different) insights about performance in a controlled laboratory situation, as well as the usage in the actual everyday life.

6. REFERENCES

[1] Barkhuus, L. and Brown, B. 2009. Unpacking the television: User practices around a changing technology. ACM Trans. Comput.-Hum. Interact. 16, 3 (Sep. 2009), 1-22.

[2] Bernhaupt, R., Obrist, M., Weiss, A., Beck, E., and Tscheligi, M. 2008. Trends in the living room and beyond: results from ethnographic studies using creative and playful probing. *Comput. Entertain.* 6, 1, Article 5 (May 2008), 23p.

[3] Bernhaupt, R. 2009. Usability and User Experience Evaluation in Non-Traditional Environments. Habilitation. University Paul Sabatier, Toulouse.

[4] Harboe, G., Massey, N., Metcalf, C., Wheatley, D., Romano, G. (2007). Perceptions of Value: The Uses of Social Television. Proc. EuroITV 2007.

[5] Harboe, G., Metcalf, C. J., Bentley, F., Tullio, J., Massey, N., and Romano, G. 2008. Ambient social TV: drawing people into a shared experience. In Proceedings CHI '08.

[6] Rice, M. and Alm, N. 2008. Designing new interfaces for digital interactive television usable by older adults. Comput. Entertain. 6, 1 (May. 2008), 1-2

[7] Tutt, D. 2008. "'Tactical' living: a situated study of teenagers' negotiations around and interactions with living room media" Environment and Planning A 40(10)

Methods for user studies of interactive TV in a commercial context

Paul Marrow
BT Innovate & Design
Orion 3/7 PP 1
Adastral Park, Ipswich, UK
+44-1473-641456
paul.marrow@bt.com

Sue Hessey
BT Innovate & Design
Orion G PP 9
Adastral Park, Ipswich, UK
+44-1473-606434
sue.hessey@bt.com

Ian Matthews
BT Innovate & Design
Orion G PP 9
Adastral Park, Ipswich, UK
+44-1473-606433
ian.r2.matthews@bt.com

ABSTRACT
Interactive TV services have attracted much attention in a commercial context because they have the potential to generate more revenue through the scheduling of services tailored to viewer's interests. However there is the problem of finding what to watch, which is why we are interested in recommendation for interactive TV. We discuss work on user identification systems, expert evaluation of existing recommender systems, and focus groups, which have provided us with information upon which to base future field trials of recommender systems on an interactive TV service. These trials will use both quantitative (indirect, large-scale) and qualitative (direct, small-scale) methods.

Categories and Subject Descriptors
H1.2 [**Models and Principles**]: User/Machine Systems – *human factors.* H3.5 [**Information Storage and Retrieval**]: Online Information Services – *commercial services.* H5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – *video*.

General Terms
Design, Experimentation, Human Factors.

Keywords
Recommendation, Personalisation, Marketing.

1. INTRODUCTION
We conducted our research as part of a project [1] studying recommender systems. Although most recommender systems have been developed for web browsers we are interested in recommendation for an interactive TV service [2]. User studies of the interaction between users and technology have attracted much attention [3], but because we are dealing with an existing commercial service, our research is unusual in that it must not disrupt the existing service, including any marketing, scheduling and business policy. In addition, we aim to ensure this will lead to an improved user experience which in turn provides a commercial benefit to the company.

2. RELATED WORK
Recommender systems are frequently deployed on interactive web browsers, e.g. [4]. Their use in interactive TV is at an earlier stage, but is attracting a great deal of attention because of the demonstrated benefits that recommendation can bring to interactive web technologies. However, at the present time interactive TV is typically a different experience from interacting with a web page, being more a social or household activity rather than an individual one, involving the viewer being seated further away from the screen, and typically not involving the username/password identification that many websites with recommendation use.

3. METHODOLOGY
Prior to carrying out user studies, we have sought to identify information from existing principles of User Interface (UI) design, as well as evaluating the UI designs of existing recommender systems (RSs), and presenting different UI designs to focus groups. While the information that we can gain in this manner is only indirect, it does show some general principles [5].

3.1 Expert Opinion – User Identification
How should users of a TV recommender system identify themselves (or not)? We identified several key points:

a) Logging on with a username and password for individual identification to the TV system is less likely to be successful than for web-based systems, due to the limited functionality of the remote control.

b) TV-viewing is often transient in nature and the system cannot identify who may have left the room when logging in manually.

c) Avatars- virtual representations of users- in the form of customizable characters may partially solve the problem, having been used extensively in games, for example. There is however issues surrounding this approach as avatars may be deemed childish, and privacy would still need to be achieved by inputting a password.

d) Recommender services can be based on "Time/date/device". Specific identification of an individual is not needed, but instead of the delivery device (e.g. family TV, personal TV, mobile device), and where appropriate of the time and date the request was made. This last option avoids additional input by the user and is therefore best suited to TV-based systems.

3.2 Evaluation of existing RSs
We also carried out two evaluations of UI designs of existing RSs [5], in the form of an expert evaluation and a focus group. The 8 recommender systems chosen for both evaluations were all web-based, reflecting their availability in comparison to TV-based

systems. So the results of this evaluation it may be argued apply more strongly to web-based recommender systems than TV-based ones, but we believe that the analysis, based as it was on well-established principles of UI design, gives some insights that are relevant to both online and TV interfaces.[1]

3.2.1 Expert evaluation

Nielsen's principles for UI design [6], a renowned reference for web usability, was used as a template for assessing the usability of the RSs, with a number of RS-specific attributes included to add relevance to the findings. Each RS was evaluated by one highly experienced UI designer.

Qualitative statements and quantitative ratings were applied to each attribute (where available and relevant) and captured in a spreadsheet, so that high and low-rated items could be identified quickly to draw out best- and worst-practice features of each RS. Then three main strengths and three main weaknesses were identified by the evaluators to summarize each RS's UI and listed in a table in qualitative form. Similar statements made about each positive and negative point across different systems were grouped until seven categories that appeared most frequently were identified. Table 1 shows these categories, some of which are general to web site design and some of which are more specific to recommender systems.

Table 1. Categories identified from the expert evaluation

General (relevant to all website design)	Recommender-system specific categories
Ease of use	User information
Appearance	Recommended items
Navigation/browse	Extra functionality
Help + explanation	

Subsequently this table was then analyzed to identify key themes across the UI design of all RS's evaluated.

3.2.2 Focus Groups

Focus groups [7] were identified as the third stage of UI evaluation to add insights from people not directly involved in the research projects to offer their opinions as potential users of the RSs. The focus group participants were company employees but were not participants in the project: due to the size of the company their responses were unlikely to be biased by knowledge of its objectives. The same RSs were presented to the attendees but additionally, one (out of two conducted) focus group was presented a "flat" (i.e. non-interactive) mock-up of the proposed interactive TV RS.

The focus group moderator led the participants through a series of open-ended questions covering general and specific topics relevant to the experience of each RS presented, and opened up the discussion to think freely about how this may be applied on the TV.

3.3 UI design evaluations: conclusion

The blend of methodologies used in this research exercise yielded results and conclusions which are beneficial in anticipating how viewers would react to recommendations delivered via TV when we carry out user studies on a pre-existing service (see [5] and Table 1).

[1] It should be emphasized that inclusion in, and the results of the study, do not represent an endorsement or criticism of the web site or company concerned.

4. FUTURE WORK

The user experience of interactive technology can be studied at many levels, whether it is on TV, web pages or with other technology. At one level objective measures of user interaction can log user actions and use these to report user behavior to analysts. At another level subjective responses by users can provide detailed information about an individual's own reactions. Other work has shown [8] how these are just two aspects of how reaction to recommender systems can be measured.

In order to carry out an analysis of user behavior of an interactive TV recommender service in a business context, we propose to split our study into two parts. The first will address the business benefits of recommendation in interactive TV, using quantitative information gathered indirectly from large numbers of users. The second will study user experience in more detail, through direct interaction with a smaller group of users across a few households. By combining the results of these studies we expect to gain both academic and commercial insights about recommender services for interactive TV.

5. ACKNOWLEDGMENTS

This work has been supported in part through the EU FP7 project IST-215006 MyMedia. We thank other members of the consortium for comments.

6. REFERENCES

[1] http://www.mymediaproject.org/default.aspx

[2] http://www.btvision.com/

[3] Sundstedt, V., Whitton, M. and Bloj, M. 2009 The whys, how tos, and pitfalls of user studies. ACM Intl. Conf. on Computer Graphics and Interactive Technologies, New Orleans. DOI= http://doi.acm.org/10.1145/1667239.1667264

[4] Dias, M.B., Locha, D., Li, M., El-Deredy, W. And Lisboa, P.L.G. 2008 The value of personalised recommender systems to e-business: a case study. In: ACM Conf. On Recommender Systems 2008, Lausanne. DOI= http://doi.acm.org/10.1145/1454008.1454054

[5] Meesters, L., Marrow, P., Matthews, I., Hessey, S., Bollen, D. and Graus, M.2009. User Design Specification. User Control Test Results Report. MyMedia Project ICT-2008-215006 Deliverable D1.4. Download from http://www.mymediaproject.org/deliverables.aspx

[6] http://www.useit.com/papers/heuristic/heuristic_list.html

[7] Finch, H. and Lewis, J. 2008 Focus Groups. Ch.7 in *Qualitative Research Practice: A Guide for Social Science Students and Researchers*, pp. 170-198. Sage, London.

[8] Meesters, L., Marrow, P., Knijnenburg, B., Bowhuis, D. and Glancy, M. 2009. End-user Recommendation Evaluation Metrics. MyMedia Project ICT-2008-215006 Deliverable D1.5. Download from http://www.mymediaproject.org/deliverables.aspx

Concept and Usability Testing Methodologies for Interactive TV

D.J. Wheatley
Motorola Applied Research Center
1295 E. Algonquin Rd.
Schaumburg, Illinois, USA
+1 847 538 1577
David.j.Wheatley@motorola.com

S. Basapur
Motorola Applied Research Center
1295 E. Algonquin Rd.
Schaumburg, Illinois, USA
Telephone number, incl. country code
sbasapur@motorola.com

ABSTRACT
The evolution of TV from a passive media consumption to a more complex interactive experience requires that greater attention be paid not only to the UI design but also the approaches to user evaluation. Furthermore, time/errors based methodological approaches traditionally adopted in a task-based productivity context may be inadequate to effectively evaluate this more social and entertainment use context. This paper describes two quite different user studies, with different objectives, carried out to evaluate a TV-based video communications system.

Categories and Subject Descriptors
H.4.3 [Communications Applications]: Computer conferencing, teleconferencing, and videoconferencing; H.1.2 [Models & Principles]: User/Machine Systems - Human factors; H.5.2 [User Interfaces]: User-centered design, Evaluation and Methodology

General Terms
Experimentation, Measurements, Human Factors, Usability.

Keywords
Video mediated communications, system usability scale, video telephony, user study, usability evaluations.

1. INTRODUCTION
Television viewing has traditionally been a passive media consumption experience with little interactivity beyond that of changing channels and more recently, DVR trick play. The introduction of more complex interactive functionality is not only changing and expanding the experience, but is creating more complex user-TV interactions which demand more design development and user evaluation. Two evaluation methodologies are described which were used in the development of a TV-based videotelephony system and which had quite different objectives. The first study set out to evaluate the user value of the concept of videotelephony via the domestic TV, including when, why and for what purposes such a system would be used and what features and functions might be required. A prototype system was developed based on the findings of this concept study and, in the second study, two variants of the system were evaluated for usability.

2. CONCEPT STUDY METHODOLOGY
The concept study set out to evaluate the user value of a TV based video telephony experience by comparing it to a PC/webcam experience and a face-to-face communications experience [1]. Equipment was set up to enable bi-directional audio/video communications between pairs of participants. In the PC and TV conditions, the equipment was set up in separate rooms so that no face-to-face communication was possible between the subjects except via the system. In the face-to-face condition (F2F) the subject pairs were seated at a table, opposite one-another. Each subject pair carried out three activities in each of the three conditions (F2F, PC and TV), these being controlled for order effects. Though these activities were purpose-driven, the aim was to create a social rather than a work related interaction. In keeping with the social/domestic usage context of the system, these activities were a Taboo game, a Charades game and a Lego building task. These were chosen in order to engender a fun and relaxed ambience consistent with the expected context of use. A further aim was also to encourage both verbal and non-verbal (visual/gestural) communication between subjects.

In the Taboo game, one subject was given the name of a major US city and using verbal and non-verbal communications, had to get the second subject to guess the name. The first subject was also given a list of words which could not be used. Three different city names were used and allocation to the conditions was controlled. In the Charades activity, one subject was given the name of a movie and using only visual communication methods, had to get the second subject to guess the name of that movie. These were also chosen to be reasonably challenging such that some creativity in visual communications might be required.

The third activity was a Lego building task. In this, the first subject was given a set of Lego blocks (but no instructions), while the second had only a picture of the final assembly. The objective was for the second subject to provide instructions to the first subject to build the assembly, using verbal and/or visual forms of communication and without showing the picture to the other person. As the aim was to provide the context for a goal oriented interaction between participant pairs, the success or failure of each activity was not relevant as a quantitative measure and hence was not recorded. At the beginning of each study condition, subjects also spent two to three minutes in general discussion in order to get accustomed to the medium of communication being evaluated. Each of the other activities then took between 1 and 5 minutes each, for a total "in condition" time of around 10 to 20 minutes. The order of presentation of the three activities was also controlled

for order effects. For each of the three activities, each subject independently completed a pre and post-study questionnaire to quantify their expectations of the communication method prior to the experience and their value judgments immediately following. Some of the expectations rated by users were comfort with the communication medium, quality of communication, privacy and security etc. They also individually completed a system usability scale [SUS] to enable quantitative comparison of the three conditions. In the TV condition alone, a special "features" questionnaire was also tendered to identify the relative value of some potential future features. Following each set of three conditions the subject pair was brought together for a semi-structured discussion about their experiences.

3. SUBJECTS

In order to ensure comfortable interactions between subject pairs, and to emulate the social context being investigated, same-gender pairs were recruited together to ensure that they knew one-another. Eighteen pairs participated in the study, with six pairs in each of the three age groups. The age categories were not chosen with the intent of being equal in range, but to increase the likelihood of age related differences being found in the subjective data by sampling from three disparate segments of the general population.

4. RESULTS

The post-test questionnaires indicated a significant effect of condition for all aspects, with TV rated significantly higher than PC, and F2F significantly higher than PC. However, there was no significant difference between the ratings for F2F and TV, consistent with the data from the SUS. In other words, the TV condition was rated very close to face-to-face communication. The features questionnaire indicated the two most highly rated items to be related to control; control over the sharing of video and control over who could use the system. Many subjects stated that the TV interaction felt more personal with a greater sense of connection with the other person because it was easier to see facial expressions and to sense feelings and emotions. Also, the head to waist view enabled body language to play a greater role, allowed subjects to be more expressive and there was a much greater sense of eye contact. However the self-view image generated some negative responses, partly due to being presented as picture-in-picture. Some subjects often looked at themselves and many characterised it "weird" to see themselves talking and the expressions they made, causing some to pay undue attention to their posture, hair, weight etc. This was readily explained by the fact that, in normal face-to-face communication, you do not see yourself so they were not used to it and most described it as a distraction. Subjects accepted the value of the self-view for camera set up but most would want to turn it off during a video call.

5. USABILITY STUDY METHODOLOGY

Based on these findings, a prototype application was developed and two different user interfaces designed. A usability study was carried out to evaluate the ease of use, learn-ability and usefulness of features offered. 24 subjects (mean 33 yrs) participated and each completed a total of 19 task sequences using each of the prototypes. They rated each prototype on the criteria of ease of use and the usefulness of the function demonstrated. The relative value of the features and functions was also rated in terms of the extent to which the presence/absence of the features might impact on a decision to purchase or acquire the system, one aim being to further test the subjective findings of the concept study. All ratings were done using six-point Likert scales. The tasks included making and receiving video calls, camera adjustments (digital pan, tilt, zoom, camera view presets, self-view on/off and camera auto-aim) and changing of video settings to make audio only calls. The task sequence within each prototype condition was fixed but the conditions were controlled for order effects. The study environment was set up to have the look and feel (including ambience and lighting) of a typical domestic sitting room and the subjects sat on a couch with a coffee table in front much as they would do at home. Prototype user interfaces were operated using a standard TV remote control at a viewing distance of about 10 feet.

6. DISCUSSION

These two user studies were undertaken at different stages of the project and used quite different methodologies because they sought to address different research questions. The concept study was conducted at a very early stage of the product concept development to identify the user value of the *concept* of TV-based video telephony and the social/contextual benefits and drawbacks relative to some existing communication methods. This was used to steer the system design in terms of features, functions and market positioning. Usability was not an element of this study, in fact at this time; the UI had not even been designed; the concept evaluation study actually guided the design team in the UI design. One key lesson learned was that we may have gained a better understanding of how subjects felt about the concept if we had used some kind of "emotional state" assessment tool. We chose not to use such a tool in the study since they are time consuming to tender and to analyse the data. There are tools such as AttrackDiff™ which address some of those issues [3]. Where similar study objectives exist, we would likely try out a tool such as this to assess the emotional experience of users. The more traditional usability study was carried out later in the development with two objectives; to confirm that the features and functions identified in the first study had been integrated in a way which was useful, and secondly to evaluate the usability of the interface design. In both cases, alternatives were evaluated to enable subjects to make comparative rather than absolute judgments.

7. REFERENCES

[1] Wheatley, D.J., and Basapur, S., 2009. A Comparative Evaluation of TV Video Telephony with Webcam and Face to Face Communication. Proc. of EuroITV. Leuven, Belgium, pp 1-8, Jun 2009.

[2] Harboe G., Massey N., Metcalf C., Wheatley D.J., & Romano G. "Perceptions of Value: The Uses of Social Television" Proc. of EuroITV 2007, Amsterdam, Netherlands, 24-25 May 2007

[3] Hassenzahl, M. (2002). The effect of perceived hedonic quality on product appealingness. International Journal of Human-Computer Interaction, 13, 479-497. http://www.attrakdiff.de/en/Home

A Pragmatic Approach to Testing Issues in a Mobile Platform that Does Not Yet Exist

The CAMMP Story

Alexandre Fleury
Aalborg University
Dpt. Of Electronic Systems
Niels Jernes vej 12
9220 Aalborg, Denmark
amf@es.aau.dk

Jakob Schou Pedersen
Aalborg University
Dpt. Of Electronic Systems
Niels Jernes vej 12
9220 Aalborg, Denmark
jsp@es.aau.dk

Lars Bo Larsen
Aalborg University
Dpt. Of Electronic Systems
Niels Jernes vej 12
9220 Aalborg, Denmark
lbl@es.aau.dk

ABSTRACT

This paper reports on user studies conducted in a project dealing with a converged mobile rich media platform. We address the questions of the selection of test methods and their application through an iterative process. We argue that a varying degree of simulation of test variables such as the environment, the service or the mobile device behavior optimizes the evaluation in terms of cost-efficiency and results quality, while providing flexibility and ease of use to the researchers conducting the evaluations.

Categories and Subject Descriptors

H.5.2 [**Information interfaces and presentation (e.g., HCI)**]: User Interfaces – *Evaluation / Methodology*

General Terms

Experimentation, Human Factors

Keywords

User studies, methods, iterative approach, mobile media

1. INTRODUCTION

The question of how to evaluate mobile services with end users has been addressed by many researchers, concerned either with defining an evaluation framework [1], selecting the methods [5], the test environment ([2]) or the degree of realism of the prototype used [6]. However it seems that no formal approach exists for conducting user studies in the context of a long lasting project dealing specifically with mobile rich media. Instead, ad-hoc methodologies seem to prevail because of various constraints such as the project organization or the amount of resources available.

1.1 Context of the Study

The work reported in this paper takes place in the context of the Danish Converged Advance Mobile Media Platforms (CAMMP) project[1], which addresses the convergence of media services with mobile technologies. In this purpose, CAMMP merges 3G mobile technologies with Internet, digital TV and radio and investigates the potential of this new infrastructure which combines traditional media and user-generated content. Eight academic and industrial partners are involved in one or more of the five work packages covering areas such as User requirements & testing (for which the work presented in this paper is conducted), Physical infrastructure or Services and content architecture.

[1] http://www.cammp.aau.dk/

2. PRAGMATIC, ITERATIVE APPROACH

CAMMP is a four-year project which follows a yearly iterative cycle. Each work package constantly and progressively develops its knowledge and/or skills in its focus area. In the authors' work package, the iterative approach to user studies is highly inspired by Nielsen's guideline concerning iterative design ([7], p. 106), which recommends to combine iterative design and evaluation. This approach has been largely used in previous software development projects and appears reliable. However in our case, the ``product'' of interest is a platform that combines various concepts, services and devices that are under investigation and/or development, and therefore not necessarily available yet. The purpose of the iterative process is then to build a pool of knowledge on issues ranging from general concepts to specific usability issues related to the platform, relying on test setups of varying degrees of simulation. At the end of the project, CAMMP will have uncovered a selection of focus areas including user expectations, design, functionality, usability, pricing, etc. Between each iteration, the issues of interest for the next iteration are mostly derived from previous results. The choice is also coordinated with the other work packages in order to synchronize the overall project development.

2.1 Combining Traditional Test Methods

The methods used in CAMMP include traditional usability evaluation ones carried out either in a laboratory, a simulated environment or outside in the field. A non-definitive list of methods relevant for the testing activities in CAMMP is reported in [8]. These methods include *Individual interviews* and *Panel discussions* (task-based, situated, in a lab, open, semi-structured, etc.), *Usability evaluations* (lab-based, in a simulated environment) and *field trials* (large scale, unsupervised).

The order in which these methods are used and the choice of methods to be applied to investigate a specific issue depend on three factors: cost-efficiency, result quality and prototype availability. With regards to the latter, since the project goal is to investigate new technologies, future services or interaction paradigms, conducting user tests requires simulating part of the setup. For instance the first iteration included a high level conceptual evaluation of mobile broadcast TV acceptability. For

this study, we used DVB-H capable Nokia phones. However, the DVB-H network infrastructure being not available in Denmark yet, a portable broadcast solution had to be specially developed. From this investigation some issues emerged, from both the participants and the facilitators. The latter then selected some of the issues to be further investigated during the next iteration. A new set of prototypes and simulating strategies was therefore required and thus developed. As depicted in Figure 1, this way of conducting test iterations leads to the discovery of new issues to be investigated in a cyclic fashion.

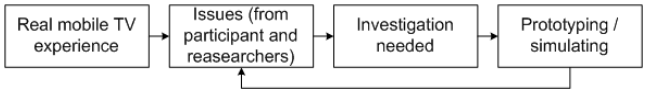


Figure 1 CAMMP's iterative test cycle

2.2 The Case of Transition Delays

The issue of transition delays while browsing through mobile television channels illustrates the pragmatic approach presented in this paper. The first test iteration included two interview-based experiments. This choice of methods for the first iteration follows common practice that favors low-cost informal methods [9] for lab evaluation of mobile products [4]. As described in details in [3], the first evaluation consisted in a face-to-face task-based interview in a busy public setup where participants were asked about the general acceptability of watching mobile television in such an environment. An existing mobile television solution was used during the test, which was evaluated by the participant in terms of first impression and general usability issues. The second test consisted of semi-structured interviews conducted in groups discussing the concepts of competition and collaboration on mobile devices.

Most participants taking part in the first evaluation as well as some participants from the second study reported that one of the main issues with the demonstrated application was the time it took the system (approximately 6-7 seconds) to change channel when users pressed the 'next channel' button. Therefore, after disseminating this finding to the other project partners, it was decided to pursue the investigation of the issue in the following iteration.

During the second test iteration (currently being conducted), a task-based evaluation setup is used to investigate this issue. In addition to identifying a threshold of acceptable delay when switching between TV channels on a mobile phone, the impact of both the environment (sitting in a quiet lab or in a simulated bus) and the type of transition (black screen with waiting icon or smooth transition) are evaluated. In later iterations, it is intended to validate these results in a setup offering an increased realism for instance with a fully functional media player tested in an unsupervised field trial.

3. CONCLUSIONS

The iterative approach to user testing discussed in this paper has so far allowed the CAMMP researchers to conduct a diversity of user experiments, investigating various aspects of the user experience with a mobile rich media platform. The work conducted in the two first test iterations involved a rather high degree of simulation and a low degree of realism. In the upcoming test iteration, the realism will be increased by conducting experiments out in the field instead of in a lab, and by introducing prototypes at a higher stage of development. Doing so will also decrease the amount of simulation required to conduct proper experiments.

4. REFERENCES

[1] Coursaris, C. K. and Kim D. J. 2006. A qualitative review of empirical mobile usability studies. In Proceedings of the Twelfth Americas Conference on Information Systems (Acapulco, Mexico, August 4-6, 2006). AMCIS 2006, 2885-2897.

[2] Duh, H. B., Tan, G. C., and Chen, V. H. 2006. Usability evaluation for mobile device: a comparison of laboratory and field tests. In Proceedings of the 8th Conference on Human-Computer interaction with Mobile Devices and Services (Helsinki, Finland, September 12 - 15, 2006). MobileHCI '06, vol. 159. ACM, New York, NY, 181-186. DOI= http://doi.acm.org/10.1145/1152215.1152254

[3] Fleury, A., Pedersen, J. S., and Larsen, L. B. 2009. Two perspectives on mobile television: Consumption in a social context and collaborative/competitive behaviors. In Proceedings of the Ninth Danish Human-Computer Interaction Research Symposium, (December 2009), 18-21.

[4] Kjeldskov, J. and Graham, C. 2003. A review of mobile hci research methods. In Human-Computer Interaction with Mobile Devices and Services, vol. 2795/2003, 317-335.

[5] Kjeldskov, J. and Stage, J. 2004. New techniques for usability evaluation of mobile systems. In International Journal of Human-Computer Studies, 60(5-6), 599-620.

[6] Lim, Y., Pangam, A., Periyasami, S., and Aneja, S. 2006. Comparative analysis of high- and low-fidelity prototypes for more valid usability evaluations of mobile devices. In Proceedings of the 4th Nordic Conference on Human-Computer interaction: Changing Roles (Oslo, Norway, October 14 - 18, 2006). A. Mørch, K. Morgan, T. Bratteteig, G. Ghosh, and D. Svanaes, Eds. NordiCHI '06, vol. 189. ACM, New York, NY, 291-300. DOI= http://doi.acm.org/10.1145/1182475.1182506

[7] Nielsen J. 1994. Usability Engineering. Morgan Kaufmann.

[8] Pedersen, J. S., Fleury, A., Sørensen, L., Heinze, E., Nicolajsen, H. W., Winbladh, J., and Wieland, J. L. 2009. State of the art literature study of methods for evaluation framework. CAMMP project deliverable D1.2.

[9] Vredenburg, K., Mao, J., Smith, P. W., and Carey, T. 2002. A survey of user-centered design practice. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Changing Our World, Changing Ourselves (Minneapolis, Minnesota, USA, April 20 - 25, 2002). CHI '02. ACM, New York, NY, 471-478. DOI= http://doi.acm.org/10.1145/503376.503460

Open Profiling of Quality as a mixed method approach to study multimodal experienced quality

Dominik Strohmeier
Institute for Media Technology
Ilmenau University of Technology
P.O. Box 10 05 65, 98684 Ilmenau
+49 (0) 3677 / 69 2671
dominik.strohmeier@tu-ilmenau.de

ABSTRACT
To reach users' overall quality acceptance is a very important aspect for the success of new technologies. Therefore, user studies following a psychoperceptual evaluation approach have been used for a long time. However, these studies neglect the dualistic nature of quality. While they are able to capture quality in its holistic excellence, methods to understand underlying quality factors are still missing. Open Profiling of Quality (OPQ) is a mixed method approach that combines quantitative and descriptive quality analysis with naïve participants. OPQ is meant as a tool to create deeper understanding of perceived multimodal quality.

Categories and Subject Descriptors
H.1.2 **[User/Machine Systems]**: *Human information processing*

General Terms
Measurement, Human Factors, Experimentation

Keywords
Open Profiling of Quality, subjective quality, quality evaluation, Quality of Experience, mixed method

1. INTRODUCTION
The optimization of quality according to users' perception has become a very important part in the development of new multimedia technologies. To provide better experience and higher quality than existing systems is a crucial point for the success of new technologies. Modern quality research needs to take into account the different aspects of quality. Beside the evaluation of quality in the actual context of use, understanding of underlying quality factors is a main goal of user-centered Quality of Experience research [1]. Recommendations on quality evaluation methods have targeted a quantitative evaluation of critical components to measure their quality excellence [8]. But quality in terms of its dualistic nature is not only related to the degree of excellence of a product. Quality is also characterized by distinctive attributes or characteristics that are possessed by products. Only few research approaches exist that try to capture the dualistic nature in subjective quality evaluation. In this paper, we present Open Profiling of Quality (OPQ) as a mixed method research approach to study perceived multimodal quality.

2. RELATED WORK
While the produced quality of multimedia application relates to the quality provided by the system under its constraints related to technical factors, perceived quality describes the users' perception of quality. Perceived quality is characterized by an impact of all levels of human perception processes [3]. Multimodal quality assessment has tried to measure the perceived quality on a quantitative basis [1]. Results of these methods present quality in terms of its hedonistic excellence. However, quality is a "*multidimensional construct of user perceptions and behaviors"* [15] and only few methods exist that offer possibilities to study and understand this construct. Bech et al.'s RaPID method [1] was a first profiling approach, but it is limited through extensive training with test participants. Focusing on naïve assessors, Jumisko-Pyykkö et al. [4] and Radun et al. [7] use interview-based methodologies to combine quantitative evaluation with an elicitation of experienced quality factors. In contrast to interviews, Lorho uses methods of sensory profiling in his Individual Vocabulary Profiling [5] approach. It allows assessors to develop and to apply their own quality attributes in the evaluation task. However, this approach has not been applied in mixed method research.

3. OPEN PROFILING OF QUALITY
OPQ [12] is a mixed method that combines evaluation of quality preferences and the elicitation of idiosyncratic experienced quality factors. It therefore uses quantitative psychoperceptual evaluation and, subsequently, an adaption of Free Choice Profiling. The goals of the method are a) to define the excellence of overall quality, b) to understand the characteristics of quality perception, c) to construct a link between preferences and quality attributes, and d) to provide a test methodology that is applicable to use with naïve assessors. OPQ does not limit assessors in the development of quality attributes to certain predefined parameters. It is *open* for assessors to elicit whatever they perceive as quality factors.

3.1 Procedure
The procedure of OPQ is described in detail in [10] and [12]. First, a psychoperceptual evaluation is conducted which follows the recommendations of the ITU [8]. Test participants rate the acceptance of overall quality and the satisfaction with overall quality. Then, a sensory profiling evaluation follows in a second task. In our research approach, we adapt the method of Free-Choice Profiling [14] which allows users to evaluate overall quality using their own vocabulary. After an introduction, test participants develop their own quality attributes, preferably adjectives, in the attribute elicitation task. In the following attribute refinement task test participants select their final vocabulary by identifying those attributes that a) are unique and b) can be described precisely by the users. At the end of the attribute refinement task, each attribute is attached to a 10cm long line labeled with 'min' and 'max' in its extremes. It results in an

individual score card which the test participants will use for stimuli evaluation. In the sensory evaluation task, test stimuli are presented one after another. The test participants mark the sensation of the attributes on their score card for each test item. 'Min' means that the attribute is not perceived at all while 'max' refers to its maximum sensation. Stimuli can be repeated several times during the evaluation.

3.2 Analysis

The application of Open Profiling of Quality results in two independent data sets. The quantitative data set can be analyzed using Analysis of Variance to get a preference ranking of the items under assessment. The sensory data can be analyzed using Generalized Procrustes Analysis (GPA). The result of GPA is a low-dimensional model of the users' perception. This model shows the separation of the test items along its principle components as well as the correlation of principle components and idiosyncratic attributes. By identifying their meaning the researcher gets deeper insight into underlying quality rationale to explain possible quality preferences in the quantitative data. The final step of OPQ is the combination of quantitative and sensory data. Using External Preference Mapping [6] allows the researcher to map preference data into the perceptual model. This finally connects the users' preferences to their quality attributes.

4. APPLICATION OF OPQ IN 3DTV RESEARCH

Open Profiling of Quality has been developed in constructive research in three studies related to 3DTV research. For detailed results of each study I refer to the respective full papers [10][11][12][13]. The first study targeted the evaluation of different audiovisual 3D presentation setups [13]. The application of OPQ has helped to see that participants were qualitatively able to differentiate variables, although no quantitative excellence of stimuli was identified. Further, different sensorial preferences as a part of any multimodal quality evaluation were identified. The second and third studies were both conducted in the context of mobile 3DTV research. The second study [11] examined influences of mono and stereo audio and visual presentation format on perceived overall quality and perceived depth. The third study finally targeted the selection of an optimum coding method for mobile 3D television and video [10]. The application of Open Profiling of Quality has helped to understand the quality rationale of autostereoscopic perception. The perceptual models of the second and third study both show that video quality is the determining component in 3D quality perception. The expected added value of depth perception only was visible when the perception of artifacts was low. These results confirm Seuntiens' 3D Quality of Experience model [9] and extend it to a hierarchical interaction of video quality and depth perception.

5. CONCLUSION

In this paper, I present Open Profiling of Quality as a mixed method approach in multimodal quality assessment. OPQ closes the shortcomings of commonly applied quantitative research methods. It combines the evaluation of the quality excellence as well as the characteristics of quality for the items under assessment and so allows capturing quality in terms of its dualistic nature. Although further work needs to study the application of OPQ in audiovisual quality research, OPQ is a valuable tool in the context of multimodal quality evaluation.

6. ACKNOWLEDGMENTS

MOBILE3DTV project has received funding from the European Community's ICT programme in the context of the Seventh Framework Programme (FP7/2007-2011) under grant agreement no 216503.

7. REFERENCES

[1] Bech, S., Hamberg, R., Nijenhuis, M., Teunissen, C., de Jong, H., Houben, P., and Pramanik, S. 1996. "The RaPID perceptual image description method (RaPID)". In Proc. SPIE. Vol. 2657, pp.317-328.

[2] Jumisko-Pyykkö, S. Strohmeier, D. "Report on research methodologies for the experiments", Technical report Project Mobile3DTV, November 2008

[3] Jumisko-Pyykkö, S., Reiter, U., and Weigel, C. 2007. „Produced quality is not perceived quality - a qualitative approach to overall audiovisual quality". In Proceedings of the 3DTV Conference.

[4] Jumisko-Pyykkö, S., Häkkinen, J., and Nyman, G. 2007. "Experienced quality factors qualitative evaluation approach to audiovisual quality". Proceedings of the IS&T/SPIE 19th Annual Symposium of Electronic Imaging, 2007

[5] Lorho, G. 2005. "Individual Vocabulary Profiling of Spatial Enhancement Systems for Stereo Headphone Reproduction". Proceedings of Audio Engineering Society 119th Convention, New York (NY), USA

[6] McEwan, J.A. 1996 "Preference Mapping for Product Optimization", in Multivariate Analysis of Data in Sensory Science, ed. Naes. T and Risvik, E., Amsterdam: Elsevier

[7] Radun, J., Leisti, T., Häkkinen, J., Ojanen, H., Olives, J.-L., Vuori, T., and Nyman, G. 2008. Content and quality: Interpretation-based estimation of image quality. ACM Trans. Appl. Percept. 4, 4, 1-15

[8] Recommendation ITU-R BT.500-11. 2002. Methodology for the Subjective Assessment of the Quality of Television Pictures, ITU Telecom. Standardization Sector of ITU.

[9] Seuntiens, P.J.H. 2006. "Visual Experience of 3D TV", PhD thesis, Eindhoven: Technische Universiteit Eindhoven

[10] Strohmeier, D. Tech, G. 2010. "Sharp, bright, three-dimensional: open profiling of quality for mobile 3DTV coding methods," in *Proc. SPIE Electronic Imaging Conf. 2010*, San Jose, California, USA

[11] Strohmeier, D., Jumisko-Pyykkö, S., Kunze, K. "New, lively, and exciting or just artificial, straining, and distracting? A sensory profiling approach to understand mobile 3D audiovisual quality," in *Proc. 4th VPQM Workshop*, Scottsdale, USA, Jan. 2010.

[12] Strohmeier, D., Jumisko-Pyykkö, S., Kunze, K. 2010. "Open Profiling of Quality: A mixed method approach to understanding multimodal quality perception," submitted to Hindawi Advances in Multimedia

[13] Strohmeier, D., Jumisko-Pyykkö, S., Reiter, U. 2010 "Profiling Experienced Quality Factors of Audiovisual 3D Perception", accepted at Qomex 2010, Trondheim, Norway.

[14] Williams, A.A., Langron, S.P. 1984. "The use of Free-choice Profiling for the Evaluation of Commercial Ports." In: Journal of the Science of Food and Agriculture 35

[15] Wu, W., Arefin, A., Rivas, R., Nahrstedt, K., Sheppard, R., and Yang, Z. 2009. Quality of experience in distributed interactive multimedia environments: toward a theoretical framework. In Proceedings of the 17th ACM International Conference on Multimedia, Beijing, China

Triangulating UX Methods for Targeting End-Users' Subjective Experiences of Media Content

Annika Wiklund-Engblom
MediaCity, Åbo Akademi University
Box 311, 65101 Vaasa, Finland
+358503223458
annika.wiklund-engblom@abo.fi

Anette Bengs
MediaCity, Åbo Akademi University
Box 311, 65101 Vaasa, Finland
+358503131952
anette.bengs@abo.fi

Susanne Sperring
MediaCity, Åbo Akademi University
Box 311, 65101 Vaasa, Finland
+35863247405
susanne.sperring@abo.fi

ABSTRACT
Our user experience research is foremost applied research targeting product development and end-user evaluation carried out in close co-operation with media content development companies. The essence of this paper is to present our exploration of methods for capturing the uniqueness of end-users' subjective experiences of media content. Both method triangulation during data collection and data triangulation during analyses are applied.

Categories and Subject Descriptors
H5.2 User Interfaces: Evaluation/Methodology

General Terms
Measurement, Design, Human Factors

Keywords
User experience, UX, triangulation, recall interview, media content development, content testing

1. INTRODUCTION
The joy-factor is a key to success for almost any media solution today. The concept of user experience (UX) emphasizes the totality of emotion, motivation, and action in a given physical and social context [1]. Users' experiences of media products and media content are affected by both product-centered aspects, such as functionality and aesthetics, as well as person-centered aspects, such as personal motivation and expectations [2]. Learning about users' subjective media experiences is a powerful tool, not to say an essential determinant for improving product design [3], which in our case implies media content creation.

Our research goals concern the understanding of end-users' media experiences. The objective is to unveil the needs, preferences, and experiences of users, viewers and consumers. One corner stone is the triangulation of methods for answering research questions related to users' attitudes, affect, actions, and reactions from media encounters. This is achieved by using a multi-method approach, in which both subjective and objective data is combined to explore the subjective experience of the end-user. The Stimulated Instant Recall (SIR) Interview is one method used to tie together other data collected during a test session. Another method is streamlining user tasks with for standardized questionnaires.

2. A MULTI-METHOD APPROACH
Earlier studies conducted at our audience research lab have confirmed that a combination of methods guarantees a broad and deep understanding of how humans experience media interactions [4]. By combining different methods (method triangulation), it is possible to study a broader range of research questions and to gain a more complete picture of end-users' experiences. It is also possible to provide stronger evidence for a conclusion through convergence and corroboration of findings (data triangulation), to increase the generalizability of the findings, and to complement any weaknesses of a single method [5, 6].

Our UX research design includes methods targeting both quantitative and qualitative data about attitudes, affect, actions, and reactions of media users. The selected methods can further be categorized into subjective and objective measurement techniques. The subjective measures include standardized questionnaires based on Likert-scales, questionnaires with open-ended questions, unstructured interviews, and various forms of structured and semi-structured interviews that are based on instant recall techniques. The objective measures include psychophysiological data collection such as monitoring heart rate, and skin conductance. Objective measures also include behavioral data recordings (e.g., eye movements, screen recordings, and recording the person within the physical environment while interacting with media content).

Figure 1 presents an outline of chosen methods and how these are distinguished regarding what they measure and what kinds of data they generate. We emphasize on method triangulation, including streamlining research methods and instruments for answering research questions in accordance with specific demands during media content development.

Methods employed for targeting emotional aspects of UX are behavioral observations (actions), psychophysiological measurement (reactions), standardized questionnaires, SIR-interviews, as well as unstructured interviews. The SIR-interviews, in which each participant is subjected to instant recall stimulated by their own replies regarding their emotions, bodily reactions, and/or behavioral actions, complement the standardized research instruments by targeting more in-depth qualities of users' emotional experiences. Stimulated recall techniques are based on an introspective research procedure for targeting the cognitive processes of research subjects. The stimulated recall triggers their memories of their thoughts during their interactions with the content [7]. For instance, eye tracking-based recall interviews

have shown to provide more verbal data than traditional think-aloud methods [8]. The validity and reliability of such eye tracking-based recall interviews has been good in HCI studies on end-user activities [9]. However, the targets of SIR-interviews and methods triangulated depend on research questions posed and the context of study, i.e. it is case dependent.

In an Attitude & Affect-Targeted Interview, for instance, questionnaire ratings are used as a recall tool during the interview. We use various techniques to highlight high and low scores, which might signal critical areas. This is a method of structuring interviews, in which a large number of questionnaire variables are down-sized into an easily manageable number. By triangulating the data gathered from questionnaires, we focus interviews in order to gain a deeper understanding of what users say about their subjective experiences.

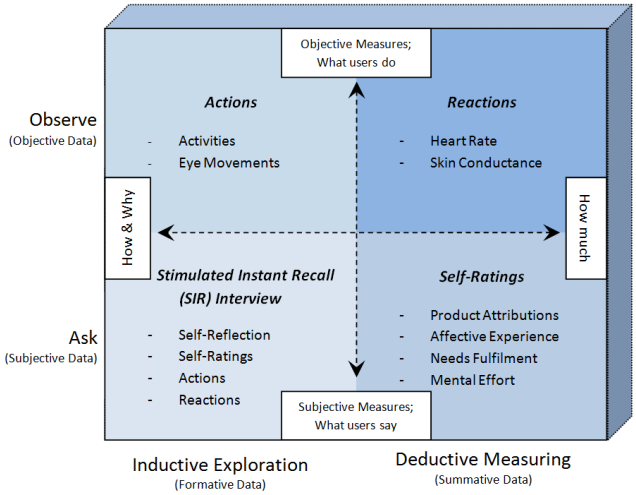


Figure 1. Method triangulation for investigating usability and user experience of media solutions, presented in relation to levels of data and research targets.

An Action & Reaction-Targeted Interview, on the other hand, is based on psychophysiological data (heart rate and skin conductance) and behavioral data (activities). Hence, the targets are what users do and how they react. Video-recordings of the user's actions and reactions during the test situation are used as a stimulus during interviews. Here the in-house developed research tool eValu8 is used, which allows a simultaneous interpretation of video-recordings/screen recordings and psychophysiological reactions. These interviews may be video-recorded in order to get as much information as possible for the analyses of stimuli of reactions and actions.

3. CONCLUSION

In the present paper, we described a multi-method approach to collect objective and subjective data for measuring users' emotional experience of media interactions. The main goal of our methods development is to guide media content development, so that the end-product answers to the needs and demands of targeted users. This is usually carried out during several phases in the development process.

The multi-method approach enables validation, triangulation and refinement of selected methods. Here, the instant recall interview technique plays a significant role. These kinds of interviews allow for method triangulation during and data collection, as well as data triangulation during analyses. It further helps us to validate findings from a wide selection of objective and subjective measurements. Our unique technique for using multiple methods and triangulating data provides a solid base for distinguishing patterns of end-users' subjective experiences, which aid the media content development.

However, we see that the applicability of existing UX methods for various content and contexts could be explored even further. The research questions raised during media content development are usually context specific. Therefore, we see the methods development as an ever continuous process in relation to targets and research questions.

4. REFERENCES

[1] Wiklund-Engblom, A., Hassenzahl, M., Bengs, A., & Sperring, S. 2009. What needs tell us about user experience. In Proceedings of the 12th INTERACT conference on Human-Computer Interaction. DOI= 10.1007/978-3-642-03658-3_71

[2] Vyas, D. and van der Veer, G. C. 2006. Rich evaluations of entertainment experience: bridging the interpretational gap. In Proceedings of the 13th European Conference on Cognitive Ergonomics: trust and Control in Complex Socio-Technical Systems (Zurich, Switzerland, September 20 - 22, 2006). ECCE '06, vol. 250. ACM, New York, NY, 137-144. DOI= http://doi.acm.org/10.1145/1274892.1274921

[3] Hassenzahl, M. 2001. The Effect of Perceived Hedonic Quality on Product Appealingness. International Journal of Human-Computer Interaction. 13 (4), 481-499.

[4] Sperring, S., & Strandvall, T. 2008. Viewers' experience of a TV quiz show. International Journal of Human-Computer Interaction. 24 (February 2008), 214-236.

[5] Brannen, J. 2005. Mixing methods: The entry of qualitative and quantitative approaches into the research process. International Journal of Social Research Methods. 8 (July 2005), 173-184.

[6] Johnson, R. B., & Onwuegbuzie, A. J. 2004. Mixed methods research: A research paradigm whose time has come. Educational Research. 33 (October 2004), 14-26.

[7] Lyle, J. 2003. Stimulated recall: a report on its use in naturalistic research. British Educational Research Journal, Vol. 29 (6).

[8] Guan, Z., Cuddihy, E., & Ramey, J. 2006. The validity of the stimulated retrospective think-aloud method as measured by eye tracking. Proceedings of the SIGCHI conference on Human Factors in computing systems. Pp. 1253-1262.

[9] Hyrskykari, A., Ovaska, S., & Tampere, M. L. 2006. Gaze path stimulation in retrospective think-aloud. Journal of Eye Movement Research. Vol. 2 (4), pp. 1-18.

Using Multiple iTV Displays for the Simultaneous Comparative Evaluation of Parallel Prototypes

Pierre-Alexandre Lapointe
User Experience Director
Yu Centrik
1654 Mont-Royal est
Montreal, QC, Canada
1 (514) 523-5862
palapointe@yucentrik.ca

Jay Vidyarthi
UX Designer / Research Coordinator
Yu Centrik
1654 Mont-Royal est
Montreal, QC, Canada
1 (514) 523-5862
jay@yucentrik.ca

Joelle Stemp
President / UX Specialist
Yu Centrik
1654 Mont-Royal est
Montreal, QC, Canada
1 (514) 523-5862
jstemp@yucentrik.ca

ABSTRACT
Best practices in iTV interface design can often be well represented in a wide range of design solutions. In an attempt to build and evaluate an optimized interactive TV interface under various constraints, we selected a combination of usability evaluation methodologies and used three high-definition television displays to obtain comparative user data on parallel prototypes simultaneously. The resulting data was instrumental in fueling subtle interaction decisions which led to a focused and optimized human-computer interface.

Categories and Subject Descriptors
D.2.2 [**Software Engineering**]: Design Tools and Techniques – *user interfaces.*
H.5.2 [**Information Interfaces / Presentation**]: User Interfaces – *evaluation/methodology, prototyping, user-centered design*

General Terms
Measurement, Design, Experimentation, Human Factors, Theory.

Keywords
iTV, User Experience, Usability Testing, Methodologies, Interface Evaluation, Human-Computer Interaction

1. INTRODUCTION
In technological design practice, scientific literature often plays the role of defining the constraints within which designers build effective human-computer interfaces. Inevitably, decisions concerning the effectiveness of contextual and subtle interaction elements which are beyond the scope of this literature must be agreed upon by project stakeholders. In 2009, brainstorms and discussion led to a collection of legitimate design concepts for the interactive TV interface of a major telecommunications provider in Canada. In such a situation, many have advocated the design of parallel prototypes to enable the direct comparison of different solutions [5, 6, 8, 10]; however, traditional qualitative and quantitative usability testing methods become laborious, repetitive, inefficient and redundant when used comparatively on multiple solutions to the same design problem [4]. As our project's time and budget were constrained, we needed to select an effective and efficient interface evaluation method to elicit user data on a number of subtly different design concepts.

2. METHODOLOGICAL SELECTION
Lacking an established usability methodology tailored to our situation, we explored the combination of related methods.

Nielsen [8] has described the benefits of designing and comparing parallel interaction modes for a single human-computer interface. The strength of this approach lies in combining the best elements of several design concepts, increasing the probability of effectiveness in the conglomerate design [6, 10]. In a similar vein, Buxton has been a proponent of comparative usability testing through the criticism of design processes which settle on a single concept too early [1]. Clearly our design concepts could be expressed through several parallel prototypes, but how could we evaluate them efficiently in a realistic usage context?

Our proposed interfaces all seemed to align with best practices; as mentioned, they differed in more subtle interaction modes. Since the resulting prototypes were all based on a similar framework, there was a clear value in considering comparative usability testing [11, 12]. However, this added even further constraint to the selection of an effective evaluation methodology: serial comparative evaluation would place a significant burden on participants' memory. Forcing participants to recall and discuss such nuanced differences would likely result in participants interacting with the interfaces under a significant cognitive load [7, 9], subtracting from the simulation of a realistic *lean-back* TV-watching experience [2, 3]. Considering this, we had no choice but to rule out these out-of-the-box methodologies; we needed a protocol which could generate firm evidence toward final design decisions within a realistic TV-watching context.

Knowing that the project budget could support 21 participants for an hour and a half each, we weighed the advantages and disadvantages of a few different methods. With a between-subjects design, we would be able to test each prototype in a unique group of 7 participants. However, with such low numbers our results would be fairly vague and inconclusive. On the other hand, a within-subjects design which presented all 21 participants with all 3 prototypes would create significantly more concrete data. Even so, in addition to the aforementioned problems with cognitive load, limited testing time would severely restrict the number of tasks and depth of discussion for each prototype. While searching for a better method, we hypothesized that the comparative evaluation of parallel prototypes might become effective through the use of corresponding parallel displays.

3. RESULTING METHODOLOGY

The resulting test procedure involved three high-definition television screens, each configured with interactive prototypes carefully designed to juxtapose subtle design elements. As a precursor to a retrospective cognitive walkthrough, the test moderator had all 21 users use parallel prototypes to perform 6 realistic TV-watching task scenarios (i.e. record a show using the EPG, choose and purchase a video-on-demand, etc.). After users commented openly, a forced choice comparison was used to determine subjective user preferences. Follow-up questions were also used to determine whether their least preferred interfaces had any specific elements which were superior to the others. Finally, a card sorting exercise was used as a manifestation of the client's desire to place an additional evaluative focus on categorization and labeling.

To build association and aid memory, tasks were initially performed with each interface on its respective screen. During the subsequent walkthrough, user attention was directed toward all three screens to enable the simultaneous consideration of multiple interfaces for each task. We hypothesized that users' familiarity with the scenarios from initial task completion combined with a simultaneous visual presentation of the prototypes would enable effortless and accurate discussion of subtle design differences.

Figure 1, The user's point of view; three high-definition screens are presented simultaneously and labeled 1, 2 and 3 to facilitate and encourage comparative discussion.

4. RESULTS AND CONCLUSION

The proposed methodology was effective in comparing and contrasting both individual features and the experiential whole of our parallel prototypes. For example, between proposed alternatives for the system's start-up interface, one model was clearly preferred. Yet in the case of video-on-demand, users identified advantages with each interface, leading to a hybrid design concept. It is also interesting to note that a wide range of user commentary and opinion significantly evolved the state of every single concept; no interface screen remained as-is.

Based on appropriately selected task scenarios and a unique testing infrastructure, we were able to compile user performance with subjective commentary and explicit preferences to draw concrete conclusions on subtle design issues. We attribute the richness of commentary elicited during the test to the fact that participants appeared to be comfortable with the approach; not only did they compare the prototype with ease, they did not seem to be intimidated by the three-screen interface display. When the testing phase was complete, we had conclusive results to fuel informed decisions on a large majority of our contentious design issues. These results have inspired a new prototype which will be subjected to a more standard usability test in 2010.

Every design project for every interface in every context has its own unique requirements. Identifying an appropriate methodology can be daunting. This case demonstrates how a creative and combinatory approach can help provide a tailored and focused methodology. Constraints inherent in our specific situation narrowed our approach to a few key methodologies. While keeping project objectives in mind, we were able to combine the most pertinent elements of relevant methodologies into an innovative hybrid approach which was effective in focusing our design efforts and moving the project forward.

5. ACKNOWLEDGEMENTS

The authors would like to thank our participants, the client organization and the EuroITV committee administration.

6. REFERENCES

[1] Buxton, B. 2007. *Sketching User Experiences: Getting the Design Right and the Right Design.* An Imprint of Elsevier. Morgan Kaufmann Publishers, Amsterdam, Netherlands.

[2] Chorianopoulous, K. 2007. Interactive TV Design That Blends Seamlessly with Everyday Life. In *Universal Access in Ambient Intelligence Environments*. Stephanidis, C. and Pieper, M., Eds. Springer, Berlin, Germany. 43-57.

[3] Griffiths, M. 2003. Internet Gambling: Issues, Concerns and Recommendations. *Cyberpsychology & Behaviour*. 6(6): 558-559

[4] Kostiainen, K., Uzun, E., Asokan, N. & Ginzboorg, P. 2007. *Framework for Comparative Usability Testing of Distributed Applications.* Nokia Research, Helsinki, Finland.

[5] Macbeth, S.A., Moroney, W.F., & Biers, D.W. 2000, Development and evaluation of symbols and icons: A comparison of the production and focus group methods, In *Proceedings of the IEA 2000/HFES 2000 Congress*, San Diego, CA, March 2000. 327-329.

[6] McGrew, J. 2001. Shortening the human computer interface design cycle: A parallel design process based on the genetic algorithm. In *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, Minneapolis, MN, October 2001. 603-606.

[7] Mullan, B. 1997. *Consuming Television*. Oxford, Blackwell Publishers.

[8] Nielsen, J. 1993. *Usability engineering*. An Imprint of Elsevier. Academic Press, San Diego, CA, 85-87.

[9] Olson, J. & Olson, G. 1990. The Growth of Cognitive Modeling in Human-Computer Interaction since GOMS. *Human-Computer Interaction*, Volume 5, 221-265.

[10] Ovaska, S. & Raiha, K.J., 1995. Parallel Design in the Classroom. In *Proceedings of the ACM CHI 95*, Denver, Colorado. ACM Press, New York, NY, 264-265. DOI= http://doi.acm.org/10.1145/223355.223666

[11] Rubin, J. 1994. *Handbook of Usability Testing.* John Wiley & Sons.

[12] Tullis, T. 2009. Tips for Usability Professionals in a Down Economy. *Journal of Usability Studies*, Volume 4, Issue 2, 60-69.

WORKSHOP 3

New Dimensions in the Assessment and Support of Quality of Experience (QoE) for Multimedia Applications)

Editors:

Dr. Shelley Buchinger (Univ. of Vienna, Austria)
Prof. Rui Jorge Lopes (Inst. de Telecomunicacoes, Portugal)
Satu Jumisko-Pyykö (Tampere Univ. of Technology, Finland),
Hans-Jürgen Zepernick (Belkinge Institute of Technology, Sweden)

Controlled vs. Uncontrolled Degradations of QoE – The Provisioning-Delivery Hysteresis in Case of Video

Thomas Zinner, Tobias Hossfeld
University of Wuerzburg,
Institute of Computer Science,
Chair of Communication Networks
{zinner|hossfeld}@informatik.uni-wuerzburg.de

Tahir Nawaz Minhas, Markus Fiedler
Blekinge Institute of Technology,
School of Computing,
SE-371 79 Karlskrona, Sweden
{tnm|mfi}@bth.se

ABSTRACT

This paper applies the recently proposed provisioning-delivery hysteresis for Quality of Experience (QoE) to the case of video. The study is based on evaluations using the Structural Similarity Metric (SSIM) for different versions of a video in terms of resolution on one hand, and suffering from different packet loss ratios on the other hand. Upon translation of the SSIM into MOS, the QoE plotted versus the effective throughput shows the predicted behaviour: a controlled quality and throughput reduction leads to a better user perceived quality than the quality degradation due to packet loss. The results clearly quantify the necessity to control quality, instead of "getting hit" in an uncontrolled way.

1. INTRODUCTION

A user of a multimedia service may choose from a variety of service and network offers. From the user point of view, the perceived cost/performance ratio is an important criterion whether to remain with or whether to leave a provider. A typical trigger for the latter, also called *churn*, are recurring quality problems for a (paid) service. For a service provider, it is thus important to understand the relationship between user perception and performance characteristics of the service provisioning through networks. In [5] a generic relationship between Quality of Experience (QoE) and Quality of Service (QoS) is investigated, categorized and formulated through various differential equations yielding exponential or logarithmic QoE-QoS relationships. In the course of this work, it became apparent that the impact of resources (such as link capacity) on QoE differs substantially from the impact of success or failures in delivery (such as packet losses). It was observed that a certain controlled reduction of the effective throughput (goodput) of a connection, e.g. through traffic shaping, affected the QoE to a much lesser extent that the uncontrolled reduction through loss, which leads to a strong reaction of TCP and increases the response times significantly. This effect is called *provisioning-delivery hysteresis* in [5] and illustrated for interactive services. Simultaneously, similar observations for streaming video traffic were made in [9]. Based on these references, our paper illustrates and discusses the provisioning-delivery hysteresis for streaming traffic. The remainder of the paper is structured as follows. In Section 2 we present the experimental setup and the estimation of the user perceived quality. Section 3 discusses the obtained results. Finally we conclude the paper in Section 4.

The work has been conducted in the project G-Lab, funded by the German Ministry of Educations and Research (Förderkennzeichen 01 BK 0800, G-Lab) and within the project "QoEMoVi" (d-nr 2006/0208) founded by the Swedish Knowledge Foundation. The authors alone are responsible for the content of the paper.

2. EXPERIMENTAL INVESTIGATION OF THE USER PERCEIVED QUALITY

This section describes the setup of the measurements we conducted and discusses the mapping between SSIM and video QoE .

2.1 Uncontrolled Distortion of a Video Clip

For the conducted measurements we used the setup depicted in Fig. 1. As operation system Debian Sid with Kernel 2.6.26-2-686 was used for all three hosts. One host, a Pentium IV equipped with a 2.4 GHz processor and 1 GB RAM, acted as video streaming server, and another, a Dual Pentium III with 2x 1.2 GHz and 512 MB RAM, acted as client. The experiments were traced using tcpdump [2]. We emulated network behavior like packet loss with a Pentium III with a 0.8 GHz processor and 256 MB RAM, located between the other hosts. As emulation software we used NETEM [6]. We investigated packet loss ratios $p_l \in [0, 5\%]$. The video sequences were transmitted using the Evalvid framework, cf. [7]. The framework provides an ap-

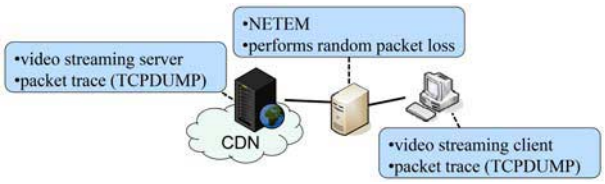


Figure 1: Measurement Setup

proach computing the received video clip of the client with the packet traces and the original video.

2.2 Controlled Distortion of a Video Clip

For streams with lower bandwidth we resized the video files with Virtual Dub [3] to a lower resolution. We encoded the clips with AutoX264 [1] and computed the average bandwidth. For investigating the impact on the user perceived quality we used a lossy upscaling to the original resolution based on the nearest neighborhood mechanism, cf. [9].

2.3 Computation of the User Perceived Quality

In order to estimate the influence of controlled and uncontrolled video distortion, we used SSIM [8] as full reference metric. As mapping function between SSIM and MOS we used the exponential fitting function $f(x) = 13.91 \cdot e^{1.715 \cdot x}$ presented in [4]. It has to be noted that the used mapping function allows MOS values $MOS \in [1,44; 3,86]$.

3. RESULTS

This section discusses the results of our investigation. Controlled reduction includes deliberate changes of resource parameters, such as resolution, image quality or frame rate, while uncontrolled reduction results from random packet losses. In this work we focus on a controlled quality reduction due to lower resolution. For our investigation we investigated three different video clips and differentiate the corresponding curves in the result figures by different colors. In [5] the provisioning-delivery hysteresis for elastic applications like TCP-based surfing was discussed. The results for this use-case are illustrated in Figure 2. The resource-related satisfaction rating function corresponds to a capacity reduction. For the case of elastic traffic protocols and applications, the throughput is adapted to network conditions. Thus, the controlled adaptation yields to still a high MOS value. On the other hand, an uncontrolled change due to e.g. packet loss might entail a quite large change of user satisfaction as displayed by the success-related satisfaction rating function. In

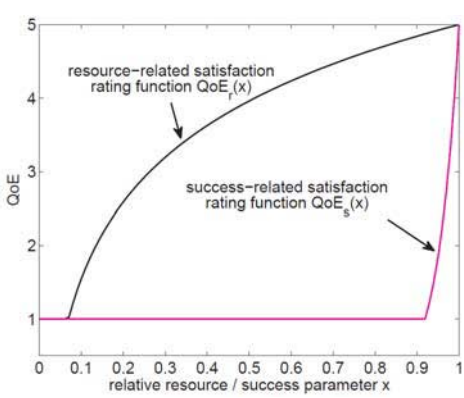


Figure 2: Provisioning-Delivery Hysteresis

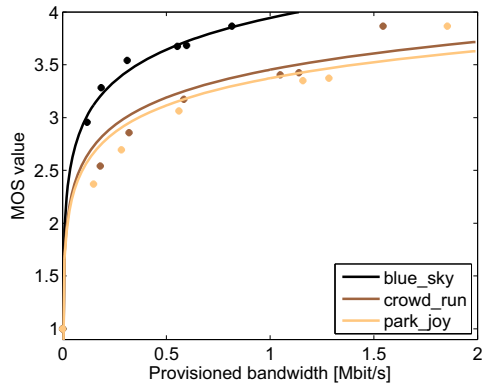


Figure 3: Provisioning Curve: MOS depending on the provisioned bandwidth

the following we discuss this behavior for video streaming.

3.1 Provisioning curves

First we shortly discuss the impact of throughput reduction on the perceived user quality. For that we investigated different resolution pairs and their required bandwidth for three different video clips. The results are depicted in Figure 3. It can be seen, that the QoE increases with the available throughput in roughly a logarithmic manner for the different video clips. Further, it can be seen, that an available bandwidth of $\approx 0,5$ Mbit/s yields a good user perception, although the required bandwidth for the best quality is much higher ($\approx 1-2$ Mbit/s).

3.2 Delivery curves

This subsection discusses the impact of uncontrollable video distortion due to packet loss on the QoE. The results are illustrated in Figure 4. The figure indicates that the already a very small packet loss $p_l < 1\%$ has a severe impact on the MOS values and thus on the Quality of Experience. For packet loss ratios $p_l > 1\%$ stalling occurs.

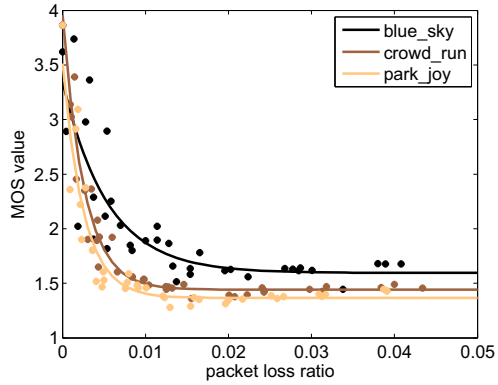


Figure 4: Delivery Curve: Impact of packet loss on the MOS

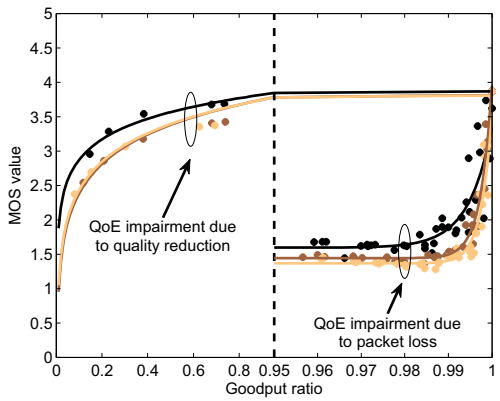


Figure 5: Provisioning-Delivery Hysteresis for video streaming

3.3 Provisioning Delivery Hysteresis for Video Streaming

Now, we combine the results for the provisioning and the delivery curves. For that we investigate the curves with respect to the effective goodput for the application. The results are plotted in Figure 5, showing MOS values as function of the goodput. We can recognise two sets of curves:

1. the upper, concave ones, emanating from variations of the resources;
2. the lower, convex ones, emanating from goodput reduction due to losses.

Further, we observe, that the impact of packet loss on the user perceived quality is much more severe than the impact of a controlled bandwidth reduction due to resolution. Fitting functions for the illustrated provisioning and delivery curves are depicted in Table 1 and Table 2.

4. CONCLUSION

The investigated behavior can be used to solve current problems in networking. For the case of mobile video streaming via CDMA systems, for instance, it is expected that a reduction of throughput can allow the use of a more robust code, which implies a reduction of delay variation, loss and the risk of outages. This may have to be paid with a reduction in image quality. However, according to the provisioning-delivery hysteresis demonstrated in this paper, such a reduction should be perceived much less critical as the frequent disturbances of the video that are more likely to appear in higher-rate scenarios, employing less robust codes.

Future work will address the quantification of the upper branch of the provisioning-delivery hysteresis for other resource-related impairments (such as frame rate, colour space, etc.) and its validation towards real users.

Table 1: Fitting for controlled quality reduction

Video	MOS	R^2
blue_sky	$0.43\ln(7653.5x)$	0.963
crowd_run	$0.62\ln(461.1x)$	0.957
park_joy	$0.60\ln(578.8x)$	0.962

Table 2: Fitting for uncontrolled quality reduction

Video	MOS	R^2
blue_sky	$1.59 + 1.49\text{e}{-80}\exp(184.4x)$	0.828
crowd_run	$1.44 + 8.14\text{e}{-167}\exp(383.4x)$	0.967
park_joy	$1.37 + 2.88\text{e}{-172}\exp(395.7x)$	0.930

Acknowledgments

The authors want to thank Hans Henjes, Michael Lang, Benjamin Kleine and Michael Denkler for their support during the course of this work. A special thank goes to Phuoc Tran-Gia for valuable comments and fruitful discussions.

5. REFERENCES

[1] Autox264. http://www.videohelp.com/tools/AutoX264.

[2] tcpdump/libcap. http://www.tcpdump.org/.

[3] Virtual dub. http://www.virtualdub.org/.

[4] Ulrich Engelke, Maulana Kusuma, Hans-Jürgen Zepernick, and Manora Caldera. Reduced-reference metric design for objective perceptual quality assessment in wireless imaging. *Signal Processing-Image Communication*.

[5] Markus Fiedler and Tobias Hoßfeld. Quality of Experience-Related Differential Equations and Provisioning-Delivery Hysteresis,. In *21st ITC Specialist Seminar on Multimedia Applications - Traffic, Performance and QoE*, Miyazaki, Japan, March 2010.

[6] The Linux Foundation. Compression project. http://www.compression.ru/video/index.htm.

[7] J. Klaue, B. Rathke, and A. Wolisz. EvalVid - A Framework for Video Transmission and Quality Evaluation. In *Proc. of 13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation*, Urbana, Illinois, USA, September 2003.

[8] Z. Wang, L. Lu, and A. Bovik. Video quality assessment based on structural distortion measurement. *Signal Processing: Image Communication*, 19(1), January 2004.

[9] Thomas Zinner, Osama Abboud, Oliver Hohlfeld, Tobias Hoßfeld, and Phuoc Tran-Gia. Towards QoE Management for Scalable Video Streaming,. In *21st ITC Specialist Seminar on Multimedia Applications - Traffic, Performance and QoE*, Miyazaki, Japan, March 2010.

Subjective Quality Evaluation of H.264 High-Definition Video Coding versus Spatial Up-Scaling and Interlacing

Yohann Pitrey, Marcus Barkowsky, Patrick Le Callet, Romuald Pépion
Image & Video Communications (IVC) Group
Institut de Recherche en Communications et Cybernétique de NANTES (IRCCyN)
École Polytechnique de l'Université de NANTES – FRANCE
{yohann.pitrey, marcus.barkowsky, patrick.lecallet, romuald.pepion}@univ–tes.fr

ABSTRACT
The upcoming High-Definition format for video display provides high-quality content, especially when displayed on adapted devices. When combined with video coding techniques such as MPEG-4 AVC/H.264, the transmission of High-Definition video content on broadcast networks becomes possible. Nonetheless, transmitting and decoding such video content is a real challenge. Therefore, intermediate formats based on lower frame resolutions or interlaced coding are still provided to address targets with limited resources. Using these formats, the final video quality depends on the postprocessing tools employed at the receiver to upsample and de-interlace these streams. In this paper, we compare the full-HD format to three possible scenarios to generate a full-HD stream from intermediate formats. We present the results of subjective tests that compare the visual quality of each scenario when using the same bitrate. The results show that using the same bitrate, the videos generated from lower-resolution formats reach similar quality compared to the full-HD videos.

General Terms
HDTV, Subjective Quality Assessment, MPEG-4 AVC/H.264

1. INTRODUCTION
Recent technological advances make transmission of High-Definition (HD) streams possible on broadcast networks, reaching a continuously increasing number of users. Combined with continuously growing network transfer capacities, this breakthrough was made possible by the advances in the video coding area. One of the most recent and most efficient is MPEG-4 AVC/H.264 [5]. This standard summarizes all the efforts in video compression from the last 20 years, and achieves good compression rates while maintaining good visual quality [3].

Currently, several HD video formats are used by service providers. The so-called full-HD format 1080p50 consists in frames of 1920×1080 pixels, displayed at 50 frames per second [1]. This format is quite demanding, because of the large size and high temporal frequency of the images to display. Therefore, intermediate formats are used to address smaller screens, networks with lower transmission rates and devices with limited processing abilities. Some formats use lower-resolution frames such as the 720p50 format (1280×720 pixels), or they use interlacing such as the 1080i50 format (1920×540 pixels per field). For displaying, the video often needs to be up-scaled and de-interlaced in order to match the full-HD native display resolution. Many techniques exist with variable complexities to perform this type of postprocessing and reduce the visual artifacts. In interlaced mode (in opposition to progressive mode), each actual frame is divided into two fields (the top and the bottom field) containing half the number of lines in the original frame. To get a full-size frame, the two fields are merged and refreshed alternatively, with half the temporal frequency of the original sequence.

In this paper, the performance of the full-HD format is compared with spatially down-scaled and/or interlaced formats, in terms of subjective visual quality. The new MPEG-4 AVC/H.264 video coding standard is used to generate streams that can be transmitted over current broadcasting channels. Using the same bitrate for each format, the impact of the coding artifacts is compared to the impact of spatial up-scaling and/or de-interlacing. Three different bitrate scenarios are proposed to study the performance of each format in low-, medium- and high-bitrate contexts. The influence of the up-scaling techniques on the visual quality is also evaluated by comparing two well-known algorithms called the Lanczos and bilinear up-scalers. Some Standard-Definition (SD) streams are included in our experiment, to show the contribution of HD over SD in a full-HD display context.

This paper is organized as follows. Section 2 presents the considered application, and the configurations tested in our experiment. Section 3 describes the subjective quality test methodology used in this experiment. Section 4 presents and discusses the experimental results. Finally, section 5 concludes the paper.

2. DESIGN OF EXPERIMENT
The purpose of our experiment is to compare different ways of generating a full-HD video stream from various input formats. In the considered scenario, only one version of the video stream would be transmitted on a network channel,

Parkjoy Crowdrun Stockholm

Figure 1: Sample of the High-Definition test sequences used for the experiment.

followed by postprocessing operations to adapt it to the requirements of a full-HD display device.

Six common video formats are studied with various frame sizes, using interlaced or progressive coding (*cf.* Table 1). To ease the reading in the following, we refer to these formats using the abbreviations defined in Table 1. Format 1 is the upcoming full-HD format. No postprocessing operation is needed to display it on a full-HD display. Currently, format 2 is used by many HD-broadcast applications. The interlacing has an impact on the quality of the displayed video, especially for scenes with high motion and high spatial detail. Format 3 is used by many service providers and it is a typical native resolution for HD-ready TV screens. The spatial up-scaling operation used to display a full-HD stream has an impact on the visual quality, due to the interpolation process. In the case of format 4, the image is only up-scaled horizontally by a ratio of 1.5. For format 5, the images further need de-interlacing combined with a similar up-scaling step. Format 6 is the current Standard Definition (SD) format. The aspect ratio is different from the HD format. The content of the output video thus appears horizontally stretched, which can affect the visual quality.

Five full-HD reference video sequences were processed. Figure 1 displays a snapshot for each one of these videos. They contain a wide variety of contents, including high motion and/or texture information. All the streams were generated from these reference sequences using interlacing and/or spatial down-scaling. The interlacing process is based on a classical convolutional technique. The down-scaling method is based on the 2D-Lanczos filter which calculates the value of a pixel from the values of the neighboring pixels in the input image, using a Mexican-hat shaped function.

The de-interlacing algorithm used is called *TDeint*. It is based on a motion-adaptive scheme to reconstruct frames from the fields in the interlaced video. It is currently known as one of the best de-interlacing methods. Two up-scaling methods are compared in our experiment. The first method is called *bilinear* up-scaling, and performs a simple 2D-linear interpolation to generate the output image. The visual performance of this up-scaling method is acceptable, while its processing complexity is comparably low. The second up-scaling method uses the *Lanczos* filter. This up-scaler is known to have good performance, at the cost of a higher complexity in terms of calculation. In the following, these two up-scaling methods are referred to using B and L.

After generating the formats reported in Table 1, each video is encoded using the reference MPEG-4 AVC/H.264 encoder

Table 1: Video formats used in the experiment.

Format abbreviation	Frame size (pixels)	Progressive / Interlace
1	1920×1080	p
2	1920×1080	i
3	1280×720	p
4	1280×1080	p
5	1280×1080	i
6	720×576	p

[4]. The encoding parameters were set to typical values, in order to get a good tradeoff between coding efficiency and computational complexity. Three bitrate scenarios were studied : 3 Mb/s, 6 Mb/s and 9 Mb/s, corresponding to low, medium and high quality of service. The non-coded full-HD video was also included in the test as a high quality reference.

For each condition tested in the experiment, three dimensions have to be specified (bitrate, resolution and up-scaling filter function). We introduce the notation of a specific configuration with the following example: 6.3.L refers to the video encoded at 6 Mb/s using format 3 (720p50) and up-scaled by the Lanczos filter.

3. SUBJECTIVE EXPERIMENT

To evaluate the visual quality of the presented scenarios, we used the Absolute Category Rating (ACR) with 11 levels methodology. The details about this methodology can be found in the IUT-T P-910 Recommendation [2]. This methodology uses a single stimulus, in which the viewer is presented one video sequence at a time. For each tested sequence, the viewer gives a score according to the level of perceived quality, located on a scale with 11 levels corresponding to adjectives ranging from *bad* to *excellent* quality.

The test was presented to 26 non-expert viewers. Their age ranged from 19 to 50, with an average of 27. Each test session took 45 minutes, starting with a visual acuity and Ishihara color plate test. According to the video contents, frame formats, up-scaling methods and bitrates included in the experiment, 145 videos of 10 seconds each were generated. After a short training, the viewer was asked to rate each video, presented in random order. They were displayed on a 40-inch diagonal full-HD reference screen (TVLogic LVM401) and the viewing distance was equal to 3 times the height of the display, according to the ITU recommendations.

4. RESULTS AND DISCUSSION

Figures 2 to 4 display the Mean Opinion Scores (MOS) obtained on each scenario. Each bar represents the average MOS of a given scenario on the five presented video sequences. The intervals of confidence are displayed as line segments at the top of each bar. They represent the variability of the opinions of the viewers on the quality scale and can be used as a first approximation for the significance of the difference between MOS. If two intervals of confidence overlap, the difference between the corresponding MOS can be considered non-significant. To provide a more precise statistical analysis, Table 2 includes the results of the student t-test, using the introduced notation. The 29 configurations tested on each video content are compared in terms of statistical difference. If two configurations are not statistically different, the corresponding cell in Table 2 contains a '·' symbol. If two configurations i and j are statistically different, it means it is possible to order them in terms of visual quality. If the MOS of configuration i is higher than the MOS of configuration j, the cell located at the i-th line and j-th column in Table 2 contains a '↑' symbol. In the opposite case (the MOS of configuration i is less than the MOS of configuration j), the cell contains a '↓' symbol.

On the three charts in Figures 2-4, the left-most score corresponds to the non-coded 1080p50 reference. As expected, this scenario obtains the highest score, as the input format does not need de-interlacing nor up-scaling, and there are no coding artifacts. The difference in MOS between the reference and the coded 1080p stream is comparably high at 3 Mb/s. It shows that the coding artifacts have a strong impact on the visual quality at this bitrate. The difference decreases in the 6 Mb/s scenario to almost 2 MOS levels, and even further for the 9 Mb/s scenario. From Table 2, we can further assess that in this scenario, the coded 1080p stream can not be differentiated from the non-coded video.

It can be observed on Figures 2 and 3 that the difference between the MOS of the SD and the HD versions is particularly high. One of the main reasons is that the spatial up-scaling does not conserve the aspect ratio, thus deforming the content. In the 3 Mb/s scenario, the average difference between the SD and the HD conditions is about 2 MOS, while in the 6 Mb/s scenario this difference is about 3 MOS. This demonstrates the advantage of HD when compared to SD in terms of visual quality, in particular for higher bitrates. For MPEG-4 AVC, the quality difference between 6 Mb/s and 9 Mb/s for the SD format was considered to be negligible during a pre-test process, therefore the 9 Mb/s SD conditions (*i.e.*: 9.6.X configurations in our notation) were not included in the experiment.

The 1080p stream coded with MPEG-4 AVC obtains a score equivalent to the scores of the up-sampled format 3 and 5 versions in the three bitrate scenarios. The impact of the coding artifacts on visual quality is thus equivalent or slightly higher than the loss in resolution from the spatial up-scaling. This constatation is particularly true in the 3 Mb/s scenario, for which the scores of format 3 are sometimes higher than the scores of format 1. Still, this difference can not be considered as significant from the results of the t-test.

The two up-scaling methods show comparable performance

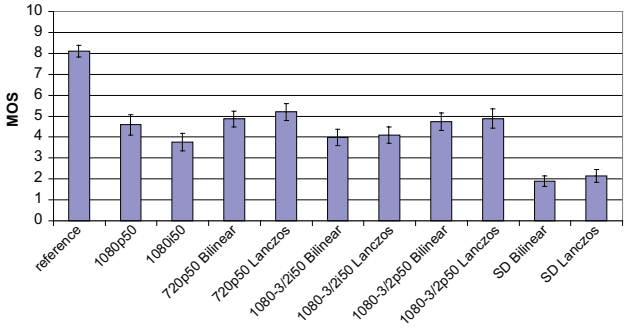


Figure 2: Mean Opinion Scores at 3Mb/s.

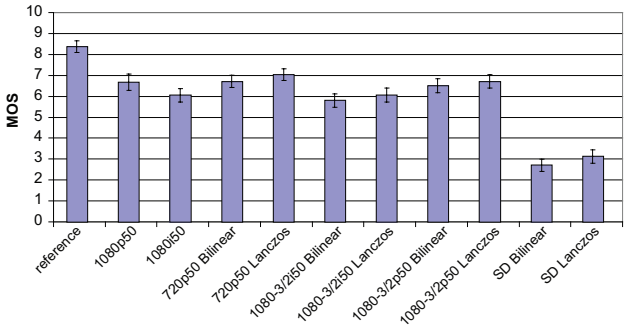


Figure 3: Mean Opinion Scores at 6Mb/s.

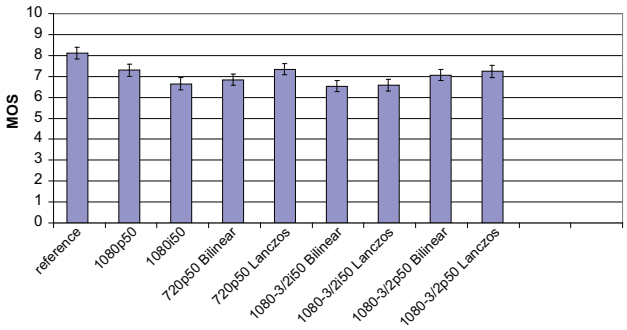


Figure 4: Mean Opinion Scores at 9Mb/s.

in the three bitrate scenarios and they are not statistically differentiable in 10 out of 11 cases. The scores of the Lanczos method are though slightly higher than the scores of the bilinear method. The bilinear method could be applied in time-critical applications or in contexts where the processing power is limited. A slight advantage can be granted to progressive streams when compared to the interlaced streams. This can be observed between format 1 and format 2, for which the quality change is always statistically significant. An interesting result is that the 1080i50 format encoded at 9 Mb/s is not statistically different from the 1080p50 encoded at 6 Mb/s. This illustrates the interest of progressive coding against interlaced coding in a medium to high bitrate context.

The results show that formats 1 and 3 are equivalent in terms of visual quality for the 3 and 6 MB/s scenarios (this is also valid in the 9 Mb/s scenario with the Lanczos up-scaling method). A possible explanation is that the down-scaling step performed before encoding the video in 720p50 from

Table 2: Results of the t-test and comparison of the MOS between the tested configurations.

	Ref	3.1	3.2	3.3.B	3.3.L	3.4.B	3.4.L	3.5.B	3.5.L	3.6.B	3.6.L	6.1	6.2	6.3.B	6.3.L	6.4.B	6.4.L	6.5.B	6.5.L	6.6.B	6.6.L	9.1	9.2	9.3.B	9.3.L	9.4.B	9.4.L	9.5.B	9.5.L
Ref		↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
3.1	↓		↑	·	·	·	·	·	·	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
3.2	↓	↓		↓	↓	·	·	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
3.3.B	↓	·	↑		·	↑	↑	·	·	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
3.3.L	↓	·	↑	·		↑	↑	·	·	↑	↑	↓	↓	↓	↓	·	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
3.4.B	↓	·	·	↓	↓		·	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
3.4.L	↓	·	·	↓	↓	·		↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
3.5.B	↓	·	↑	·	·	↑	↑		·	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
3.5.L	↓	·	↑	·	·	↑	↑	·		↑	↑	↓	↓	↓	↓	↓	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
3.6.B	↓	↓	↓	↓	↓	↓	↓	↓	↓		·	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
3.6.L	↓	↓	↓	↓	↓	↓	↓	↓	↓	·		↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
6.1	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑		↑	·	·	↑	↑	·	·	↑	↑	↓	·	·	↓	·	·	↓	↓
6.2	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↓		↓	↓	·	·	·	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
6.3.B	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	↑		·	↑	↑	·	·	↑	↑	↓	·	·	↓	·	·	↓	↓
6.3.L	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	↑	·		↑	↑	↑	·	↑	↑	↓	·	·	↓	·	·	·	↓
6.4.B	↓	↑	↑	↑	·	↑	↑	↑	↑	↑	↑	↓	·	↓	↓		·	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
6.4.L	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↓	·	↓	↓	·		·	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓
6.5.B	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	·	·	↓	↑	·		·	↑	↑	↓	·	↓	↓	·	·	↓	↓
6.5.L	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	↑	·	·	↑	↑	·		↑	↑	↓	·	·	↓	·	·	↓	↓
6.6.B	↓	↓	↓	↓	↓	↓	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓		·	↓	↓	↓	↓	↓	↓	↓	↓
6.6.L	↓	↓	↓	↓	↓	↓	↓	↓	↓	↑	↑	↓	↓	↓	↓	↓	↓	↓	↓	·		↓	↓	↓	↓	↓	↓	↓	↓
9.1	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑		↑	↑	·	↑	↑	·	·
9.2	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	↑	·	·	↑	↑	·	·	↑	↑	↓		·	↓	·	·	↓	↓
9.3.B	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	↑	·	·	↑	↑	↑	·	↑	↑	↓	·		↓	·	·	·	↓
9.3.L	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	↑	↑		↑	↑	·	·
9.4.B	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	↑	·	·	↑	↑	·	·	↑	↑	↓	·	·	↓		·	↓	↓
9.4.L	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	↑	·	·	↑	↑	·	·	↑	↑	↓	·	·	↓	·		↓	↓
9.5.B	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	↑	↑	↑	↑	↑	↑	·	↑	·	·	↑	↑		·
9.5.L	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	·	↑	↑	·	↑	↑	·	

1080p50 acts as a low-pass filter on the original video. Small irregularities and noise from the original sequence can thus be smoothed in the down-scaled version, which in some cases eases the encoding. This also holds for the receiver as the up-scaling process might counteract the blocking artifacts.

In our experiment it was observed that the up-scaled video can have a quality that is very close to the 1080p50 video when using a high quality up-scaler such as the bilinear or the Lanczos filter. Thus a transmission in 720p50, 1280×1080p and 1280×1080i formats may be considered. This might be advantageous for the sender or the receiver. Due to the smaller resolution, the computational complexity of the encoding and decoding is reduced. Since the down-scaling and up-scaling operations do usually not require an equivalent amount of computation, the overall complexity is reduced. Thus, on the sender side computational resources might be reallocated in order to evaluate more sophisticated video coding options. On the receiver side, the up-scaling is usually performed by hardware, *e.g.* the graphics card. As a result the reduced computational complexity in the decoder might lead to a longer battery life in the case of a portable device. A disadvantage of decoding a lower resolution video might be the increased memory bandwidth requirement resulting from this two-step approach in which an intermediate image has to be stored.

5. CONCLUSION

In this paper, we used the results of subjective tests to compare the performance of various High-Definition video formats coded with MPEG-4 AVC/H.264. Spatial up-scaling and de-interlacing were performed to display a full-HD stream independently from the encoded format. Using both graphical and statistical analysis, the results show that with equivalent bitrate, the difference in quality between the streams coded directly in full-HD and the interlaced and down-scaled streams is mostly not visually significant.

6. REFERENCES

[1] ITU-R BT.709-5 Recommendation. Parameter values for the HDTV standards for production and international programme exchange. 2004.

[2] ITU-T P.910 Rec. Subjective video quality assessment methods for multimedia appl. 1996.

[3] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi. Video Coding with H.264 / AVC: Tools, Performance & Complexity. *IEEE Circuits & Systems Magazine*, 2004.

[4] Joint Video Team. MPEG-4 AVC/H.264 reference encoder software. *http://iphome.hhi.de/suehring/tml/*.

[5] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra. Overview of the H.264/AVC video coding standard. *IEEE TCSVT*, 13(7):560–576, July 2003.

YoMo: A YouTube Application Comfort Monitoring Tool

Barbara Staehle, Matthias Hirth, Rastin Pries, Florian Wamser, Dirk Staehle
University of Würzburg, Institute of Computer Science, Würzburg, Germany
{bstaehle,hirth,pries,wamser,dstaehle}@informatik.uni-wuerzburg.de

ABSTRACT

Out of the large number of multimedia content sharing platforms YouTube is the most popular one. This is reflected by the large number of studies which focus on analyzing YouTube characteristics. Techniques for quantifying the instantaneous YouTube QoE and predicting an imminent QoE degradation have in contrast never been proposed. The latter task is even more important if network management actions shall be carried out to avoid a YouTube QoE degradation. In this work we describe YoMo, a tool which constantly monitors the YouTube application comfort. This measure quantifies the application operation condition and allows a QoE prediction. Experiments show that YoMo is able to exactly anticipate an upcoming YouTube QoE degradation.

1. INTRODUCTION

According to the 2009 Cisco Visual Networking Index [3], 30% of all customer Internet traffic consists of downloading or streaming videos. According to Cisco Systems, this share will increase to over 60% by the year 2013. For the case of the U.S., roughly half of this Internet video traffic is due to user-generated content whereof again roughly the half is due to YouTube. Consequently, many authors try to find reasons for the success of YouTube. The most prominent example of this category of studies is the work of Cha et al. [1] who analyze the distribution and evolution of the YouTube video popularity and user behavior. Cheng et al. [2] compare the characteristics of YouTube videos and the structure of the underlying social network. Gill et al. [7] examine YouTube usage patterns, file properties, and transfer behaviors.

In contrast, the QoE of YouTube has to the best of our knowledge not yet been in the focus of research studies. The large number of QoE models for video streaming [6] are not applicable, as they assume UDP as transport layer protocol. In this case, delay, jitter, and packet loss are directly related to the video QoE as they cause artifacts or missing frames. YouTube videos, in contrast, are transported via HTTP over TCP. No packets are lost or delivered out of order, and the only quality degradation which may be caused by the transmission, is a stalling of the video. For this case, the approach of Gustafsson et al. [8] would be applicable which allows to derive the QoE from video parameters, the packet loss rate, and the number of buffering events after the video has been terminated.

The YouTube player has no feedback loop which adapts the video playback to the network conditions. Our goal is therefore to use the YouTube QoE as an input for a network management tool which allows to maintain the user QoE. Therefore, we need to know the user satisfaction in real time. Moreover, a management tool needs a prediction of the QoE, i.e. it needs to be notified if the YouTube player is about to stall in order to avoid this. Dalal et al. [4] already proposed a QoE prediction mechanism UDP video streaming. It however uses lost and retransmitted packets only, a method which does not work in case of TCP streaming.

In this work we introduce the YouTube monitoring tool YoMo. It constantly monitors the YouTube application comfort (AC) which is simply the amount of playtime buffered by the YouTube player. In general, AC characterizes how well the application is doing. This allows YoMo to derive a simple binary YouTube QoE in real time which is either "good" if the video plays and "bad" if the video stalls. Additionally, YoMo is able to predict a QoE degradation, namely the stalling of the video. In an earlier study [5] on QoE-based radio resource management, we showed that a cooperation of YoMo and a bandwidth shaping tool allows a continuous YouTube playback in a congested wireless mesh network.

The lack of literature on YouTube QoE is the reason why we think that YoMo is interesting for more areas than network monitoring. Therefore, we discuss the idea of using the AC to predict the QoE and its implementation by YoMo in the following. Section 2 gives an overview on the core idea and implementation details. YoMo's functionality is evaluated in Section 3. Section 4 summarizes the contributions of this paper and gives an outlook to future work.

2. MONITORING THE YOUTUBE AC

To be able to monitor the YouTube AC, YoMo has to fulfill several tasks: Firstly, it has to detect that a YouTube flow exists which has to be monitored. Secondly, it has to collect as much information as possible about the YouTube

This work was funded by the Federal Ministry of Education and Research of the Federal Republic of Germany (Förderkennzeichen 01 BK 0800, GLab). The authors alone are responsible for the content of the paper.

flow and thirdly, it has to monitor the YouTube AC. To make our approach more easy to understand, we first of all analyze the technology behind YouTube in Section 2.1, before we introduce the main ideas of YoMo and their implementation in Section 2.2. How the the amount of buffered playtime is estimated is described in detail in Section 2.3.

2.1 The Technology Behind YouTube

The YouTube player is a proprietary Flash application which concurrently plays a Flash video (FLV) file and downloads it via HTTP. At the beginning of this so-called pseudo streaming, the client fills an internal buffer and starts the video playback as soon as a minimum buffer level, γ, is reached. During the time of simultaneous playback and downloading, the buffer grows as long as the download bandwidth is larger than the video rate and shrinks otherwise. If the buffer runs empty, the video stalls and the YouTube player state changes from "playing" to "buffering". This state is hidden to the normal user, but can be retrieved from the YouTube API by JavaScript or ActionScript.

Each YouTube video is encoded as an FLV file which is a container format for media files developed by Adobe Systems. An FLV file encapsulates synchronized audio and video streams. The header starts with an FLV signature and contains information about the tags in the body of the file. The tags encapsulate the data from the streams and contain information on their payload. This information includes the payload type, the length of the payload, and the time to which the tag payload applies. FLV files may also contain metadata encapsulated in a tag with a script data payload. The available properties depend on the software used for the FLV encoding and may include the duration of the video, the audio and video rate, and the file size.

2.2 The Main YoMo Functionality

The YouTube player opens a new TCP connection each time it downloads a new FLV file or if the user jumps to another time in the video. Each FLV file has a header with the FLV signature, the beginning of a new YouTube video flow is hence marked by this signature. YoMo runs at the client and parses all incoming TCP flows in order to detect this signature. Once a flow containing FLV data is recognized, the data is continuously parsed in order to retrieve the available meta information from the FLV file. Detecting the YouTube flow is thus easily done. The AC monitoring task is more complex and will be explained in the following.

The YouTube AC is defined as the buffer status of the YouTube player. This is simply the time, β, the player can continue playing if the connection to the server is interrupted. Fig. 1 shows β as the difference between the currently available playtime T and the current time of the video t. YoMo constantly computes and visualizes β in a GUI and checks whether β falls below an In such a situation, like the one depicted in Fig. 1, the QoE is still good, as the video is playing, but the AC is bad, as $\beta < \beta_a$ and the video is about to stall soon. Hence, YoMo predicts an upcoming stalling and has to notifiy a network management tool or decrease the video bandwidth in order to avoid this.

2.3 Estimating the Buffered Playtime

YoMo computes the buffered playtime as $\beta = T - t$. It decodes the FLV tags in real time, and hence exactly knows the currently available playtime T which is the time stamp

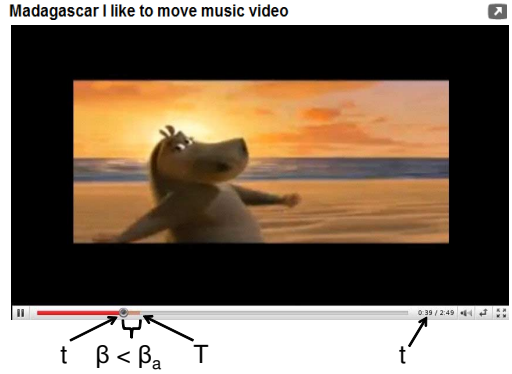


Figure 1: The YoMo Parameters

of the last completely downloaded tag. Intuitively, t could easily be calculated as the time difference between the actual time and the time when the player starts to play the video. During our measurements we found that this is not as easy as assumed. The reason for this is that the playback of a YouTube video does not start immediately after the player has loaded, but only after an amount γ of bytes has been downloaded. In [5], we show results from experiments with different videos and different connection speeds. The experiments reveal that γ is varying between 50 and 300 kB and is independent of the connection speed but is different for each of the 10 considered videos. We analyzed the coefficient of correlation between γ and different video characteristics including information about the frame types of the original H264 file embedded in the FLV tags, but were not able to find a clear correlation which allows to derive γ from the properties of the displayed video.

It is hence not possible to calculate the amount of time which lies between the time when the user issues the request for the video and the time when the video actually starts to play. We therefore implemented two different methods for calculating t which we discuss in the following. *Method 1* uses the assumption that the video starts to play as soon as the first FLV tag is completely downloaded. Clearly, this introduces a small error in the calculation of β which decreases however with an increasing connection speed. *Method 2* stands for the way of obtaining t from the YouTube player API which can be accessed by scripting languages only. In order to make YoMo applicable for the use in productive environments, it has to work with the original YouTube web page which can not be modified. It is also unrealistic to redirect all YouTube traffic to a dedicated web page where scripts for YoMo are running. Hence, YoMo uses a Firefox plugin which runs a JavaScript that retrieves t from the YouTube player. The plugin additionally sends the actual value of t to YoMo.

3. THE YOMO ACCURACY

YoMo and the Firefox plugin may be downloaded from the G-Lab website[1]. In the remainder of this section, we investigate how exactly YoMo can predict a QoE degradation. For this purpose, a client is connected to the Internet via a proxy which is able to modify the connection speed. The proxy may also interrupt the connection and thereby cause a video to stall. The client does not access the original

[1]http://www.german-lab.de/go/yomo

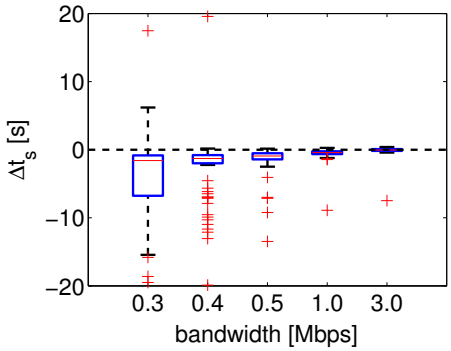


Figure 2: Stall Time Estimation Error, Method 1

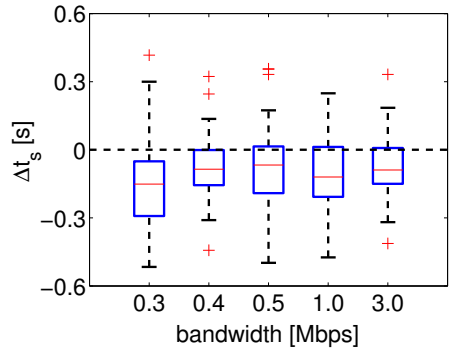


Figure 3: Stall Time Estimation Error, Method 2

YouTube side, but runs a measurement web page which embeds the YouTube player. This allows to dump the YouTube player state and thereby to get the exact stalling time. The client additionally runs YoMo which logs the estimated stall time which we consider to be the time when $\beta \leq 0.5$ sec. This is due to an experiment with 100 randomly chosen videos where observed that $\beta = 0$ sec is a sufficient but not a necessary condition for a stalling video as many videos already stall if $\beta \approx 0.5$ sec.

In Fig. 2 and Fig. 3, we depict the estimation error Δt_s between the time when YoMo considers the video to stall and the video actually stalls for Method 1 and 2 respectively. For each considered bandwidth, a box depicts the inter quartile range of the estimation errors and whiskers which are 1.5 times longer than the interquartile range. Values beyond this range are shown by red crosses. Let's discuss Fig. 2 which represents the estimation accuracy of Method 1, first. It shows that the error decreases with an increasing bandwidth. This is just a logical consequence of neglecting the time required for downloading γ, which gets smaller if the Internet connection is fast. While this method is thus sufficiently accurate for a broadband Internet access, it results in YoMo estimating the video to stall up to 20 seconds earlier as it actually did in the case of a slower connection.

The results for the experiment with estimation Method 2, shown in Fig. 3, in contrast visualize an bandwidth-independent error. Moreover does YoMo estimate the video on average to stall only roughly 0.1 sec earlier than it actually did, which is a significant improvement over Method 1. In most cases, YoMo underestimates the remaining play time, i.e. predicts the time of stalling earlier than it actually happened. The maximal estimation error in this direction is 0.5 sec. In some cases, YoMo overestimated the remain play time with a maximal error below 0.5 sec. Taking the inherent error of our assumption that a video stalls if $\beta < 0.5$ into account, these results demonstrate that YoMo, with Method 2 for the buffer estimation, is working as intended. In [5] we are moreover able to show that this accuracy is suitable for a QoE guaranteeing radio resource management.

4. CONCLUSION AND OUTLOOK

Application comfort monitoring presents an approach to monitor the usage of applications, their quality requirements and the experienced application comfort at the client. AC monitoring allows a QoE prediction which is very valuable input for network management tools. In the scope of YouTube, or more generally, Flash video streaming over TCP, the AC is measured as the buffered playtime. YoMo consists of a Java application and a Firefox plug-in. This plug-in monitors the state, in particular the current playtime, of the Flash player. A packet sniffer detects new Flash video transfers, extracts the videos metadata, and monitors the available playtime. Both components together allow to determine exactly the buffered playtime.

We were able to demonstrate that YoMo is able to accurately estimate the time when the YouTube player is stalling. YoMo is lightweight and easy to install while it provides valuable information to an ISP. If users run YoMo, both parties may greatly benefit as the provider gets information for free which it can use for improving the user QoE. Our future work will therefore be dedicated to examining the suitability of YoMo for QoE-based network management in various scenarios more closely.

5. REFERENCES

[1] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn, and S. Moon. I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System. In *SIGCOMM'07*, San Diego, CA, USA, October 2007.

[2] X. Cheng, C. Dale, and J. Liu. Statistics and Social Network of YouTube Videos. In *IWQoS'08*, Enschede, The Netherlands, June 2008.

[3] Cisco Systems Inc. Cisco Visual Networking Index - Forecast and Methodology, 2008-2013. White Paper, June 2009.

[4] A. C. Dalal, D. Musicant, J. Olson, B. McMenamy, S. Benzaid, B. Kazez, and E. Bolan. Predicting User-Perceived Quality Ratings from Streaming Media Data. In *ICC'07*, Glasgow, Scotland, UK, June 2007.

[5] B. Staehle, M. Hirth, F. Wamser, R. Pries, and D. Staehle, "YoMo: A YouTube Application Comfort Monitoring Tool," University of Würzburg, Tech. Rep. 467, March 2010.

[6] U. Engelke and H. J. Zepernick. Perceptual-based Quality Metrics for Image and Video Services: A Survey. In *NGI'07*, Trondheim, Norway, May 2007.

[7] P. Gill, M. Arlitt, Z. Li, and A. Mahanti. YouTube Traffic Characterization: A View From the Edge. In *SIGCOMM'07*, San Diego, CA, USA, October 2007.

[8] J. Gustafsson, G. Heikkila, and M. Pettersson. Measuring Multimedia Quality in Mobile Networks with an Objective Parametric Model. In *ICIP'08*, San Diego, CA, USA, December 2008.

Using metadata in video quality assessment based on the structural similarity (SSIM) index metric

Tiago Gonçalves
Instituto Universitário de Lisboa (ISCTE-IUL) – Instituto de Telecomunicações
Lisboa, Portugal
a24548@iscte.pt

Rui J. Lopes
Instituto Universitário de Lisboa (ISCTE-IUL) – Instituto de Telecomunicações
Lisboa, Portugal
Rui.Lopes@iscte.pt

Paulo Nunes
Instituto Universitário de Lisboa (ISCTE-IUL) – Instituto de Telecomunicações
Lisboa, Portugal
Paulo.Nunes@iscte.pt

ABSTRACT

In this paper, we propose the use of metadata to define the perceptual weight of semantically meaningful regions in video frames for video quality assessment. The proposed approach is evaluated using the structural similarity (SSIM) index metric and a comparison is made with both the unweighted approach and the luminance-based weighted approach. Experimental results demonstrate that the proposed approach leads to significantly different mean SSIM values relatively to the other methods under comparison, suggesting a better correlation with the human perceived visual quality.

Categories and Subject Descriptors

E.4 [**Coding and Information Theory**]: *Data compaction and compression*;
H.1.2 [**Information Systems**]: User/Machine Systems - *Human information processing.*

General Terms

Algorithms, Measurement, Experimentation, Human Factors, Verification.

Keywords

Video Quality Assessment, Visual Metadata, Structural Similarity (SSIM), Quality of Experience (QoE), EvalVid

1. INTRODUCTION

As more and more multimedia content is being delivered through best-effort IP networks, the video quality assessment methodologies become more relevant. In addition to the video encoding process and the terminal characteristics, the IP network also introduces impairments that should be considered.

In this context, the ultimate goal of quality assessment is to emulate the human's perceived quality of the 'consumed' multimedia signals, also known as Quality of Experience (QoE) evaluation.

Subjective QoE evaluation, which is, therefore, concerned with the end-user experience, is usually accomplished by measuring the Mean Opinion Score (MOS) [1]. MOS is obtained from a large group of individuals that is asked to watch a set of videos and to rate their quality continuously in real-time, using a 1–5 scale. The subjective QoE assessment must follow a precise methodology and use a well-controlled environment [7], which makes the process very expensive and cumbersome.

On the other hand, the objective video quality assessment goal is to obtain a MOS, not from individuals' opinions, but from measurable characteristics of the multimedia signals or from network measurements, such as average packet loss rate or sustainable bit rate. Peak Signal-to-Noise Ratio (PSNR), and the related quantity Mean Squared Error (MSE), are the most widely and well-known used Full Reference (FR) quality metrics used for image and video quality assessment. Nevertheless, PSNR has a limited range of validity as it measures essentially the sample-wise distortion between a reference signal and its 'impaired' version. The Structural Similarity (SSIM) index follows a different approach for video quality assessment [8], as it is based on the idea that the human visual perception is highly adapted for extracting structural information from a scene. Structures of the objects in a scene are independent of the influence of the luminance and contrast. Thus, luminance, contrast and structure are the separated components that are measured and compared. The SSIM indexing metric uses an 8×8 pixels sliding window approach, where the sliding windows moves pixel-by-pixel from the top-left corner to the bottom-right corner of the image. The overall quality value is defined as the average of the quality map, also known as the Mean SSIM (MSSIM) index.

A straightforward extension of the SSIM metric for video data, by simply averaging its value over entire frames and over time does not take into account that neither all regions in a visual scene nor all frames are equally important in terms of human visual perception. Therefore, the MSSIM averaging computed for each video frame should be modified to a weighted average, given the different importance of the various frame regions to the human observers. In addition, the same principle should be applied over time. In [9], the authors follow a simple approach of weighting differently the various sliding windows for each video frame according to their average luminance value. It is recommended to increase the weight of the windows with higher luminance, and weighting each frame in the video sequence according to its motion activity.

Other aspects are, however, involved in the perceptual relevance of the various scene regions, besides the local average luminance and motion activity values, which could be considered in weight assignment. While PSNR and SSIM are only based on the video signal, it is plausible that the semantic information has an important role on human perception of content. Therefore, in this paper, the proposed approach for quality evaluation is to use semantically meaningful metadata to describe the subjectively relevant regions in each video frame and to establish a weight mapping function according to the subjective relevance of each region. This approach constitutes a potential improvement for SSIM-based quality evaluation, but can also be used with other FR video quality metrics.

2. WEIGHTED MSSIM VIDEO QUALITY ASSESSMENT

The Structural Similarity (SSIM) index between signals **x** and **y** is

$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)},$$

where μ_x and μ_y are, respectively, the **x** and **y** averages, σ_x^2 and σ_y^2 their corresponding variance**,** σ_{xy} is the covariance between **x** and **y**, $C_1 = (K_1 . L)^2$ and $C_2 = (K_2 . L)^2$, where L is the pixel dynamic range (255 in our case), and K_1 and K_2 are two constants for stabilization purposes (in this paper $K_1 = 0.01$ and $K_2 = 0.03$, which are the typically recommend values [8]).

Notice that SSIM equals 1 when the images that are being compared are equal, and is less than 1 when they are not equal.

Since the SSIM is obtained for 8×8 windows, computing the MSSIM metrics involves averaging over these windows. Equally averaging all the windows on a single frame is the simplest way of computing MSSIM. Nevertheless, images are meaningful to humans, both in terms of pixels values, activity and relevant semantics, suggesting that a different weighting should be used.

2.1 Weights Based on Luminance

The approach proposed in [9] is based on the assumption that dark regions do not attract attention, therefore should be weighted with smaller values. Being μ_x the mean luminance for a given window and w_{ij} the weight of the *j*-th window in the *i*-th frame, the local weighting is advised in [9] to be adjusted as

$$w_{ij} = \begin{cases} 0 & \mu_x \le 40 \\ \dfrac{\mu_x - 40}{10} & 40 < \mu_x \le 50 \\ 1 & \mu_x > 50 \end{cases},$$

giving lower importance to low luminance.

2.2 Weights Based on Metadata

The alternative video quality evaluation approach that is proposed in this paper is to associate metadata information with the video in order that the importance of the video objects is defined *a priori*. How the metadata is obtained is not in the scope of this work.

The required elements for applying this method are: *i*) the semantic description of the sequence; *ii*) the definition of the frame regions associated with the semantic content; and *iii*) a mapping function of the descriptions in weights, for every region.

In this paper we have manually assigned weights to the regions that subjectively seemed to capture more visual attention. The regions were simply cataloged into three groups: *i*) high-significant regions (weight of 1); *ii*) medium-significant regions (weight of 0,5); and *iii*) low-significant regions (weight of 0,1). Figure 1 illustrates the regions used for a sample video frame and their associated weights. This particular video sequence consists of an aircraft slowly moving from left to right, but not traveling the entire image. The aircraft region is considered the most significant, followed by its boundaries and the control tower, which are half-significant. The rest of the scene almost does not capture visual attention, therefore is considered insignificant.

This technique can be used in any scenarios where the three previously elements can be defined.

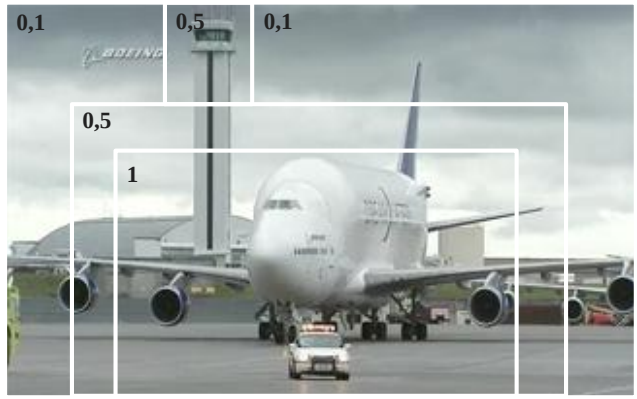


Figure 1. Example of subjectively relevant video frame regions with corresponding weights.

3. ARCHITECTURE AND TEST METHODOLOGY

3.1 EvalVid

The core architecture used for video transmission quality evaluation is an extension to the EvalVid framework. EvalVid is a toolset for video quality evaluation transmitted over real or simulated communication networks [3][4] using a FR scheme. The transmission description is accomplished by two packet traces, one from the sender and another from the receiver.

Two main extensions were made to EvalVid. The first intended to change the error model of the network simulation, from a bit error rate model to a Packet Loss Rate (PLR) model, which is more relevant in our assessment. The second main extension was to add MSSIM measurement to the platform (since only PSNR and computed MOS from PSNR were originally implemented), and to support different weighting methods: equal, luminance-based, and metadata-based.

3.2 Test Sequences and Conditions

For the tests in this paper, we have used 3 video sequences from the LIVE Wireless Video Quality Assessment Database [5] representing airport hangar activity: vid1_hp.yuv (*Sequence* 1), vid6_jp.yuv (*Sequence* 2) and vid12_jp.yuv (*Sequence* 3). Each video has 768×480 pixels with 300 frames at 25 fps.

We simulated the transmission of the encoded video sequences (VBR H.264 coded with a target bit rate of 1 Mbit/s and a bit rate tolerance of 100 kbit/s) with five different uniform PLR conditions: 0%, 1%, 5%, 10% and 20% (50 simulations each), for each video sequence. The target bit rate was chosen in order to focus the video quality assessment on the network impairments due to packet losses and not on the video coding impairments.

Then, the equal weight, luminance-based weight and metadata-based weight methods were applied. For the metadata semantic descriptions of the *Sequence* 2 and 3, the same criteria illustrated in Figure 1 for *Sequence* 1 was used.

4. TEST RESULTS

Figures 2-4 show the results of the MSSIM measurements for each video sequence. Each group of columns refers to a different PLR, being each column associated with a weighting strategy (equal, luminance-based and metadata-based). The results are the average of MSSIM over the 50 simulation runs for each strategy

and the error bars are defined for a 95 % confidence interval, with a standard normal distribution.

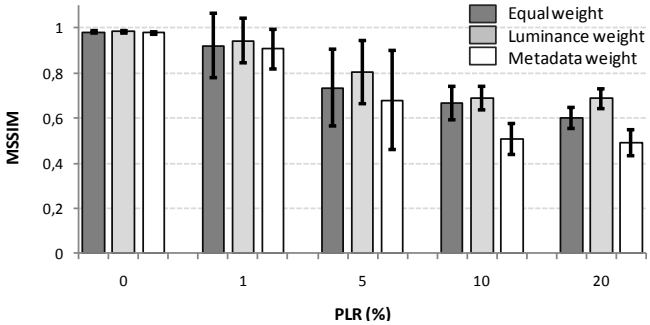


Figure 2. *Sequence* 1 MSSIM for each PLR and weighting strategy.

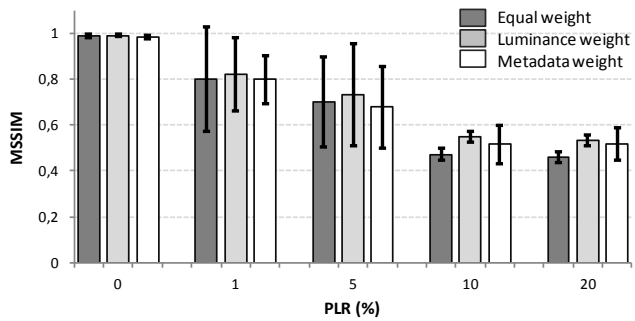


Figure 3. *Sequence* 2 MSSIM for each PLR and weighting strategy.

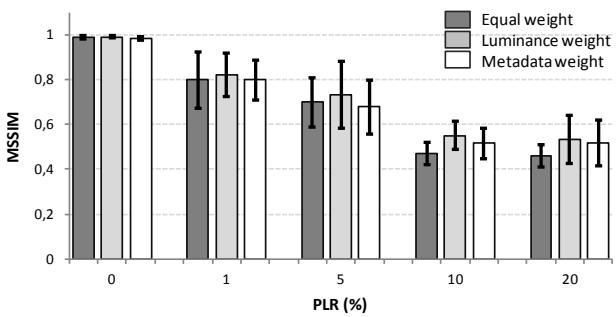


Figure 4. *Sequence* 3 MSSIM for each PLR and weighting strategy.

For all test sequences (Figure 2-4) the MSSIM values are similar for a PLR of 0%, with the average and the confidence interval not differing too much. These values although very close, are not equal to 1 because the SSIM index has been measured between the original reference video sequence and the coded video sequence. Notice that for a PLR 0% only coding impairments are taken into account, while for remain PLRs measure the network impairments are also considered.

Figure 2 shows that, the luminance-based approach always increases the MSSIM value, while the metadata-based approach always decreases it. Figures 3 and 4 do not follow this trend. Although the luminance-based approach values are always higher than the metadata-based approach, the latter is not always lower than the equal average case, namely in the worst-case transmission conditions (PLR of 10% and 20%).

5. DISCUSSIONS AND CONCLUSION

We propose the use of metadata to enhance video quality assessment based on the SSIM index metric by balancing the relevance (weight) of the various regions of a video frame. Preliminary experimental results shown that the proposed approach significantly changes the obtained MSSIM values in the presence of packet losses indicating that this strategy deserves further investigation, namely, how the MSSIM values obtained with the proposed approach correlate with MOS values obtained through subjective testing. Effective non ad-hoc ways to define the weight mapping function and the computational cost assessment are other possible topics to be explored. The weighting over time is another issue that should be studied, possibly taking into account end-user visual memory perceptual characteristics such as sensitivity to motion, and cinematographic effects like fade in/out. Future work will include also the extension of this approach to Reduced Reference (RR) metrics.

6. REFERENCES

[1] Huynh-Thu, Q. and Ghanbari, M. 2008. Scope of validity of PSNR in image/video quality assessment. Electronics Letters 44, 800–801.

[2] ITU- R, 2000. Rec. ITU-R BT 500-10 - Methodology for the subjective assessment of the quality of the television pictures.

[3] Klaue, J., Rathke, B. and Wolisz, A. 2003. EvalVid - A Framework for Video Transmission and Quality Evaluation. In Proc. of the 13th Int. Conf. of Modelling Techniques and Tools for Computer Performance Evaluation (Urbana, IL, USA, Sep. 2003).

[4] Lie, A., Klaue, J. 2008. EvalVid-RA: trace driven simulation of rate adaptive MPEG-4 VBR video. Multimedia Systems, 14, 33-50.

[5] Moorthy, A. K. and Bovik, A. C. 2009. LIVE Wireless Video Quality Assessment Database. http://live.ece.utexas.edu/research/quality/live_wireless_video.html

[6] Moorthy, A. K. and Bovik, A. C. (To appear). Wireless video quality assessment: A study of subjective scores and objective algorithms. IEEE Trans. on CSVT.

[7] Simone, F. *et al*. 2009. Subjective assessment of H.264/AVC video sequences transmitted over a noisy channel. Proc. of the 1st Int. Workshop on Quality of Multimedia Experience (San Diego, CA, USA, July 2009).

[8] Wang, Z. *et al*. 2004. Image quality assessment: From error visibility to structural similarity. IEEE Trans. on Image Proc., 13, (4), 600-612.

[9] Wang, Z., Lu L. and Bovik A. C. 2004. Video quality assessment based on structural distortion measurement. Signal Processing: Image Communication, special issue on “Objective video quality metrics”, 19 (2), 121-132.

[10] Wang, Z., Simoncelli, E. P. and Bovik, A. C. 2003. Multi-scale structural similarity for image quality assessment. Invited Paper, IEEE Asilomar Conf. on Signals, Systems and Computers (Pacific Grove, CA, USA, Nov. 2003).

Test persons for subjective video quality testing: Experts or non-experts?

Matej Nezveda, Shelley Buchinger, Werner Robitza,
Ewald Hotop, Patrik Hummelbrunner and Helmut Hlavacs
Department of Distributed and Multimedia Systems
University of Vienna
{matej.nezveda, shelley.buchinger, werner.robitza,
ewald.hotop, patrik.hummelbrunner and helmut.hlavacs}@univie.ac.at

Vittorio Baroncini, Cristina Delogu
Fondazione Ugo Bordoni
vittorio@fub.it, cristina@fub.it

ABSTRACT
Some time ago it has been understood that subjective quality assessment is crucial to determine the multimedia quality of distorted or compressed data. Generally, researchers follow the ITU recommendations to carry out their user tests. There, it is specified that at least fifteen non-experts need to act as test-persons. It seems that in many opportunities such a large number of test persons is not required to obtain a significant result. In this paper, we investigated if a smaller number of experts could be used instead of the large group of non-experts by repeating already performed experiments with the new setting.

Tests results reveal that for an adequate choice of the content, the rating scores obtained by non-experts can be approximated very precisely by ratings performed with a fewer number of experts. For an arbitrary selection of the content, the followed approach seems to be useful to detect and varify trends. Furthermore it has been shown that a certain number of experts needs to be used. Mean opinion scores of several non-experts can not be approximated by ratings performed by only one expert.

Categories and Subject Descriptors
H.1.2 [**User/Machine Systems**]: Human Factors

Keywords
Subjective quality assessment, video quality

1. INTRODUCTION
Quality of Experience has been defined by the ITU-T Study group 12 and pre-published in 2008 in the ITU-T Recommendation G.1080 [6] to be "Quality of experience (QoE) is the overall acceptability of an application or service, as perceived subjectively by the end user." This means that the user plays a crucial role when aiming at assessing Quality of Experience. In contrary to computational methods, user tests have the disadvantage that they comprise a statistical error. It is not possible to obtain the same result at each test run. In order to provide comparable results, methods on how to assess subjective audio, video and multimedia quality have been defined some time ago [4, 5, 2, 3].

One of the properties that have been defined in these recommendations consists in involving at least 15 non-experts to act as test persons. The effort and costs of such user tests are high and might exceed the needs and the budget of some investigations. There might be some or even several occasions where a smaller number of test users would be sufficient or where the large number of non-experts could be replaced by a small number of video quality experts. In fact, data analysis carried out in [10] revealed that the values of only ten non-experts are sufficient in several cases to obtain a significant and reliable result.

In this paper we investigate if the large number of non-experts that is required by [4] could be replaced by a smaller number of non experts. For that purpose we repeated the user tests published in [9, 8] with six experts only in a similar setting following the remaining specifications indicated in [4]. More details on the experimental set up is provided in Section 2, test results are described and discussed in Section 3 and conclusions are drawn in Section 4.

2. EXPERIMENT
We concluded a Single Stimulus Continuous Quality Evaluation (SSCQE) experiment based on [4], in which each video was shown only once to each expert observer. All observers have experience in the field of video quality. The goal of this experiment was to assess whether a set of experts could achieve the same results as a larger group of non-experts.

The videos that were used for playback originated from the LIVE database [9, 8]. It provides a set of 150 videos and the

corresponding scores that are based on 38 valid subjects. The LIVE database features ten different reference videos from which distorted sequences have been created. These distortions include IP-based and wireless distortions, visible as fading and ghosting, as well as MPEG-2 and h.264 compression errors, which lead to blocking and other artefacts. We decided to use only a subset of the database, the IP and wireless distorted videos. This is because in the original tests presented in [9], the experiments had been divided into two sessions so as not to fatigue the users. Also, all video impairments were created to span the overall bandwidth of possible scores, so there is no need to show all videos. Hence, a number of 80 videos has been selected for our investigation. All LIVE videos are in a medium resolution format down sampled from HD sources (768x432 pixels) and have a frame rate of either 25 or 50fps. The videos are 10 seconds long, except for one which only lasts for 8.68 seconds. Their content is selected so as to span a wide range of possible combinations of spatial and temporal complexity.

2.1 Rating methodology

Rating was performed using a continuous scale from 0-100, but with Absolute Categorial Rating (ACR) labels, based on the five-point scale {"Excellent", "Good", "Fair", "Poor", "Bad"}. The LIVE database featured one reference video per sequence, so the Hidden Reference Removal (HR) method was used to compute Difference Mean Opinion Scores (DMOS) between the reference video and its distorted equivalents [7]. As we considered only "experts" as test subjects, no training procedure was included. Each subject viewed all of the available videos in a randomized order.

2.2 Test Equipment

Videos were played back by using the *QualiView* software. It allows for capturing the score after each video. Scores are saved to a text file and then evaluated. For presentation, a consumer LCD monitor (manufactured by AOC) was used, calibrated to standard settings. Viewing conditions were set up to be as neutral as possible, with a neutral background used for the QualiView software. The experiment was conducted in a room indirectly lit by a 6500°K lamp.

2.3 DMOS Scores

Since we tried to repeat the experiments conducted for the LIVE database, only with a small number of experts, we also used the same procedure to calculate the opinion scores for each combination of video and test subject as described in [8]: First, scores for every distorted video were subtracted from the score assigned to the corresponding reference video, which was then removed from the evaluation. These difference scores were then converted into Z-values. The obtained Z-values were re-scaled to the range $[0, 100]$. Finally, for each video sequence the average score over all observers has been computed.

3. RESULTS

Since the DMOS scores computed in this experiment have been calculated following the guidelines presented in [8], the rating error caused by using only six expert as test persons instead of at least 15 non-experts is represented by the absolute value of the difference between obtained ratings:

$$E_i = \left| DMOS_i^{LIVE} - DMOS_i \right|$$

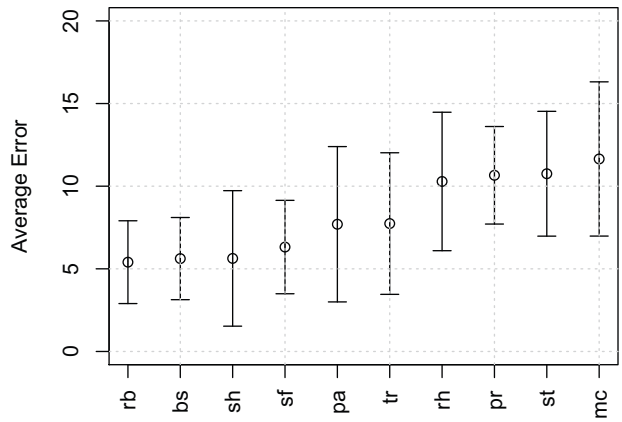


Figure 1: Average error with CI 0.95

for each video i.

In average, the computed rating error was 7.75 out of 100 scoring points with a standard deviation of 5.19. When mapping the interval from 0 to 100 to the five (ACR) categories {"Excellent", "Good", "Fair", "Poor", "Bad"}, each category is represented by an interval of a width of 20 scoring points. This means that the error in average was less than half of the width of such an interval. It has to be mentioned that the results of each two studies even when using a large number of non-experts might always return an error of a certain amount. Hence, it can be argued that the approach of using a small number of non-experts as test users reveals to be useful and precise enough for preliminary subjective tests, giving some insights and indications on trends. The result might even be improved when the experts exercise some training before conducting the test. When observing collected data in more detail, it can be observed that DMOS of some videos correlate better to the $DMOS^{LIVE}$ than others.

In Figure 1, the average rating error and respective confidence intervals are depicted for each video sequence. The sequences of the LIVE database are denoted by corresponding acronym: (rb) river bed, (bs) blue sky, (sh) shields, (sf) sunflower, (pa) pedestrian area, (tr) tractor, (rh) rush hour, (pr) park run, (st) station, (mc) mobile & calendar.

It can be observed that four sequences have an average at around 6 rating scores, then a small jump to around 7.7 scoring points with considerable standard deviation can be observed followed by average errors higher than 10 scoring points. When we started to analyze the video content in order to understand this behavior, we observed that the best correlation of DMOS scores is obtained for contents where the region of interest (ROI) as it is defined in [1] can be easily identified. For example, in "sun flower" the attention is clearly focused on a bee. In "river bed", the eye is attracted by the waves. In the sequence "rush hour" instead the focus could be put on the fountain in the background, on a car just in front of the fountain or even on a car in the bottom left. This means that if the ROI is evident, scores obtained by expert viewing are very similar to the ones obtained by a

V	ROIs	P1	P2	P3	P4	P5	P6
rb	1	9.56	15.15	7.84	9.54	7.50	9.43
bs	1	10.87	9.30	10.22	6.27	8.07	10.78
sh	1	6.99	7.57	17.17	6.82	5.69	10.73
sf	1	9.45	9.05	8.54	9.98	20.05	5.02
pa	>2	14.57	13.63	7.00	11.21	10.78	2.33
tr	2	7.38	4.58	9.13	7.69	11.67	28.51
rh	>3	11.38	17.07	9.39	15.66	7.89	11.49
pr	1	9.88	11.28	15.41	10.79	11.09	10.80
st	>2	18.36	9.11	8.97	11.93	13.88	10.89
mc	>2	7.28	13.99	18.54	9.52	5.77	20.22
avg		10.57	11.07	11.22	9.94	10.24	12.02

Table 1: Absolute error for person and video

V	ROIs	P1	P2	P3	P4	P5	P6
rb	1	6.08	8.99	7.07	6.85	6.52	3.59
bs	1	9.56	6.83	5.95	5.03	4.64	5.03
sh	1	4.23	5.90	7.25	5.04	5.35	8.72
sf	1	9.81	4.94	6.72	9.83	10.01	2.41
pa	>2	14.31	8.98	5.22	8.49	9.86	2.39
tr	2	5.30	5.00	1.03	5.89	8.87	6.57
rh	>3	8.89	13.01	8.04	9.24	4.51	4.32
pr	1	9.68	7.37	6.39	4.68	11.07	8.34
st	>2	9.83	3.59	8.18	8.72	8.42	6.20
mc	>2	7.03	10.73	8.98	10.16	4.85	9.62
avg		8.47	7.54	6.48	7.39	7.41	5.72

Table 2: Standard deviation for Table 1

larger number of non-experts. On the contrary, if the ROI is less clear, scoring results seem to correlate less. The only exception is given by the sequence "park run". The ROI should clearly be a running person, but still the rating error is significant. This can be explained by the fact that the presented park run seems to be artificial because it is shown in slow motion. Some of the experts might have detected the slow motion initially as an impairment and the mistake might be understood only after watching the reference sequence.

A more detailed view on the test results is given by the Table 1 and Table 2. For every combination of video and observer, absolute difference errors between the LIVE database values and our scores were calculated. Then, these errors were averaged for each video sequence. The errors can be seen in Table 1, their standard deviation is given in Table 2. The second column contains the number of possible ROIs where an error might have had an impact on the opinion score. It has to be noted that these results do not match with Figure 1 since the values have been calculated differently.

Since no average over the ratings performed by the test persons has been computed for Table 1 it represents the error that would have been obtained when only one expert would have been used for comparison. Under this light, the result of each test person is given by each column. It can be observed that around half of the sequences show an error of more than 10 rating points. In average, the rating error lies between 9.94 and 12.02 points. These values might be too high for a reliable approximation of subjective scores. Therefore it does not seem to be adequate to use the judgement of only one expert.

Table 2 shows the respective standard deviations. Similarly to Table 1 it seems that the standard deviations are too high when using the rating scores of only one expert.

4. CONCLUSION

In this paper it has been investigated if subjective user tests, as they are recommended to be designed by the ITU, could be simplified. Our test methodology involved a smaller number of experts as test persons instead of the recommended minimum of 15 non-experts. For videos containing only one region of interest at a time, ratings performed by six experts return a good approximation of the scores obtained from a larger number of test users. For generic content types, the results seem to be sufficiently precise to investigate general trends or to conduct some preliminary studies.

Additionally, it has been shown that the scores of only one expert cannot be used to represent the opinion of a larger number of non-experts.

5. REFERENCES

[1] U. Engelke and H.-J. Zepernick. Optimal region-of-interest based visual quality assessment. In *Proceedings of the SPIE'09, Human Vision and Electronic Imaging XIV* , volume 7240, pages 724011–724011–12, 2009.

[2] ITU-R. A guide to ITU-R Recommendations for Subjective Assessment of Sound Quality. ITU-R Recommendations BS.1283-1, 2003.

[3] ITU-R. Methodology for the Subjective Assessment of Video Quality in Multimedia Applications. ITU-T Recommendations BS.1283-1, 2007.

[4] ITU-R BT.500-11. Methodology for the subjective assessment of the quality of television pictures. ITU-R BT.500-11, 2002.

[5] ITU-T P.910. Subjective video quality assessment methods for multimedia applications. ITU-T P.910, ITU, 1999.

[6] ITU-T: Study Group 12. Quality of Experience Requirements for IPTV Services. ITU-T Recommendations G.1080, ITU, 2008.

[7] M. Pinson and S. Wolf. Comparing subjective video quality testing methodologies. In *Proc. SPIE - Visual Communications and Image Processing*, 2003.

[8] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack. Study of subjective and objective quality assessment of video. *IEEE Transactions on Image Processing*, 2009.

[9] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack. A subjective study to evaluate video quality assessment algorithms. *SPIE Proceedings Human Vision and Electronic Imaging*, 2010.

[10] S. Winkler. On the properties of subjective ratings in video quality experiments. In *International Workshop on Quality of Multimedia Experience, QoMEx.*, 2009.

Towards User-driven Adaptation of H.264/SVC Streams

Jordi Ortiz Murillo
jordi.ortiz@um.es

Michael Ransburg
mransbur@itec.uni-klu.ac.at

Eduardo Martínez Graciá
edumart@um.es

Michael Sablatschan
msablats@itec.uni-klu.ac.at

Antonio F. Gómez Skarmeta
skarmeta@um.es

Hermann Hellwagner
hellwagn@itec.uni-klu.ac.at

Keywords

H.264/SVC, Adaptation, TFRC, QoE, User-driven, Streaming

ABSTRACT

H.264/SVC enables runtime-efficient scalability in the spatial, temporal and fidelity dimension. Existing adaptation mechanisms facilitate this to automatically adapt the H.264/ SVC stream to the current usage environment without any user interaction. This paper argues that the Quality of Experience (QoE) of the end user can be enhanced by enabling him to manually adjust the adaptation if he wishes to do so. An approach which enables this is presented and evaluated. It is shown that by facilitating this approach an increased QoE is provided compared to automatic adaptation approaches. Finally, future work indicates the next steps in order to implement this approach.

1. INTRODUCTION

H.264/SVC [1][2] is a scalable video codec, which introduces scalability mechanisms in three different scalability dimensions: spatial resolution, temporal resolution and fidelity. The scalability of the encoded video bit-stream is achieved by a layered approach. A scalable H.264/SVC bit-stream comprises an H.264/AVC-conformant base layer, which represents video at the lowest quality, and one or more enhancement layers, which can be used to refine the video quality in one or more of the above mentioned scalability dimensions. The adaptation of an H.264/SVC bit-stream is performed by simply truncating these enhancement layers from the initial bit stream and can therefore be implemented very efficiently anywhere along the delivery chain. H.264/SVC thus fits the requirements of video streaming in heterogeneous usage environments and is therefore at the core of this research field, which currently focuses on pervasive adaptation. That is, the H.264/SVC stream is automatically adapted based on usage environment descriptions (e.g., client capabilities or network condition) without any user interaction. The aim is to shield the user from the adaptation process and to simply provide him with the best QoE possible given the current usage environment [3][4][5][6].

In this paper we propose a mechanism that makes it possible for the user to manually adjust the automatic adaptation process. We argue that the automatic adaptation may not always be optimal and that such a manual adjustment can help to optimize the user's QoE.

Section 2 briefly introduces the state of the art and consequently motivates our work by describing some application scenarios. Section 3 describes our approach in detail. Section 4 provides an evaluation of our approach. Section 5 gives an outlook to future work and concludes this paper

2. STATE OF THE ART AND MOTIVATION

Automatic adaptation relies on knowledge of 1) the current usage environment, e.g., the current network condition or the client capabilities and 2) content characteristics, e.g., the genre. For example, if an adaptation node is aware of the currently available bandwidth, it can select the optimal enhancement layer configuration to utilize it. If this adaptation node is additionally aware of, e.g., the display resolution of the client, it can restrict the allowed enhancement layer configurations to only these which do not degrade the spatial resolution of the video below the display resolution of the client. Further optimization of the enhancement layer configuration can be performed based on content characteristics, e.g., it is not ideal to reduce the temporal resolution of a fast-moving action movie. Taking the optimal adaptation decision based on all available usage environment descriptions and content characteristics can be formulated as a mathematical optimization problem as described in [7].

Given very complete information on usage environment and content characteristics, the automatic adaptation can provide a good QoE. However, we argue that even if such complete usage environment and content characteristics information is available, which is usually not the case, the user itself must be taken into consideration in order to increase the QoE. In the following we list some cases where automatic adaptation may benefit from manual adjustments by the user:

- Incomplete usage environment information. For example, YouTube still lets the user select the spatial

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
EuroITV2010, June 9th-11th 2010, Tampere, FINLAND.
Copyright 2010 ACM 978-1-60558-31-5/06/10 ...$10.00.

resolution of the chosen video.

- Incomplete content characteristics. In particular in live streams the content characteristics are usually not available.

- Interest of the user in a specific content. This can often change dynamically, e.g., during a news report.

- If a user's mobile data contract is limited to a certain data amount per month, the user will want to decide how much bandwidth to utilize for a video stream based on, e.g., his availability for the rest of the month, since he may be out of country.

- The viewing environment of the user, e.g., if the user wants to showcase a video streaming service to a friend, the user will most likely choose a higher quality.

It must be noted that some of the above cases might be expressed as *user preferences*, *user characteristics* or *natural environment characteristics* as a part of the usage environment description [8]. However, we believe that these are too multifaceted (as the cases above show) to be predefined. Additional challenges, such as privacy concerns, would also arise in this context, particularly when the adaptation decision is taken outside of the user's premises.

3. OUR APPROACH

Our approach, as introduced above, is not meant to replace automatic adaptation, but rather to enhance it by enabling the user to steer it. We therefore introduce two steering parameters:

1. Layer drop priority.

2. Minimum number of enhancement layers for each scalability dimension.

The layer drop priority is expressed as an ordered relation of the scalability dimensions D (for spatial resolution), Q (for fidelity) and T (for temporal resolution). This way the client is able to specify the requirements through the dropping priority of the scalability dimensions. Optionally, minimum values can be set for each of the scalability dimensions. Note that the *priority id* which is defined as a header field in H.264/SVC [1] has similar semantics, but cannot be used for expressing user preferences, since it is bound to the content.

Besides our steering parameters, we require a usage environment constraint to trigger the adaptation. Note that below we focus on the available bandwidth, however this could in theory be replaced by any other constraint which may trigger an adaptation.

The algorithm for the DQT selection, based on the steering parameters, is defined using pseudo code as follows:

Listing 1: Algorithm Pseudo-Code

```
Sort scalability dimensions according to drop priority,
from highest to lowest;

while (available bandwidth <
    bitrate of current enhancement layer selection) {
 for each scalability dimension {
  if (num layers of scalability dimension >
     min layers selected by the client) {
   Start filtering the currently highest layer
     from this scalability dimension;
   Stop filtering layers from scalability
     dimensions with higher drop priority;
   SelectionFound = true;
   End for loop ;
  }
 }
 if (SelectionFound == false) {
  Perform best effort adaptation;
 }
}
```

Obviously it would make no sense to adapt the bitstream if there is enough bandwidth available. Thus, the algorithm is only triggered if the available bandwidth is smaller than the maximum bit rate of the H.264/SVC stream. In this case layers have to be dropped in order to decrease the bit rate, as indicated by the steering parameters. This is done until a configuration of enhancement layers is found which has a bit rate smaller than or equal to the available bandwidth.

If no suitable enhancement layer configuration can be found, best effort adaptation is performed, selecting the base layer in the worst case.

4. EVALUATION

In order to evaluate our approach we extended the NS-2 simulator [9] to include RTCP feedback and TFRC calculation. Additionally, the different layer selection approaches presented below were implemented. The simulation setup consists of two nodes, representing a streaming server and a streaming client connected by a link. The bidirectional wire is configured to offer a total bandwidth of 2.9 Mb/s while the H.264/SVC video is encoded with an average bit rate of 2.89 Mb/s.

TCP Friendly Rate Control (TFRC) [10] is used to calculate the available bandwidth T for a certain session as defined in Equation 1.This function gives an upper limit to the bandwidth available in terms of bytes per second. The inputs to this functions are the packet size s, the Round Trip Time (RTT) r, the loss event rate p and the TCP retransmission timeout T_{RTO}. The implementation uses a mean of all packet sizes for s and a weighted moving average of the fraction lost for p, as shown in (1).

$$T = \frac{s}{r\sqrt{\frac{2p}{3}} + t_{RTO}(3\sqrt{\frac{3p}{8}})p(1+32p^2)} \quad (1)$$

Having calculated the available bandwidth, the DQT selection algorithm is triggered. Congestion is simulated after four seconds, as can be seen in the Figures 1(b), 2(b), 3(b) and 4(b). There is a delay until the actual adaptation begins, which can be explained with the weighted mean of the fraction lost p in Equation 1.

Another possibility to trigger the DQT selection is to manually change the available bandwidth, referred to as user triggered adaptation. This situation was simulated within the interval between seconds 14 and 18.

Four different approaches are compared in the evaluation. The first approach represents a best effort adaptation which selects the layer combination fitting best into the available

bandwidth. The second approach sets minimal values for each scalability dimension D, Q and T, but no layer drop priority. The third approach supports layer drop priority but no minimal values. Finally the fourth approach corresponds to the one proposed in this paper, including minimal values and layer drop priority.

The video used in our simulations is the City MPEG reference video sequence with 2 spatial, 2 quality and 4 temporal enhancement values. The highest layer corresponds to 4CIF spatial resolution at 30 FPS with just one quality enhancement. For both lower spatial resolutions (CIF and QCIF) a second quality enhancement was encoded.

For the evaluation the layer drop priority has been set to $D > Q > T$, the minimum values chosen were D=1, Q=1, T=2. These minimum values correspond to CIF size and a frame rate of 7.5 FPS.

Figure 1 shows the results using the first approach, i.e. best effort adaptation. Figure 2 represents the second approach with minimal values for the scalability dimensions and Figure 3 corresponds to the third approach including the layer drop priority. Finally in Figure 4 one can see the results achieved by our proposed approach using minimal values and layer drop priority combined. Each figure consists of two diagrams. The left one shows the selection of DQT values by the approaches, while the right diagram depicts the available bandwidth which triggered the selection, as well as the corresponding selected bit rate.

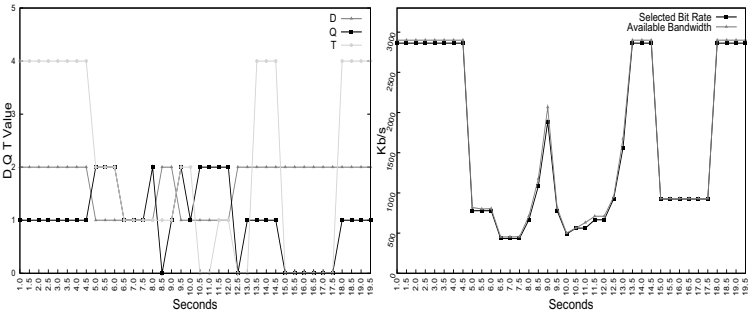
In the interval between the seconds 4 and 5.5 one can observe that the approaches with layer drop priority $D > Q > T$ (Figure 3 and Figure 4) try to keep the temporal value as high as possible. When looking at seconds 10 to 11 in Figure 3(a) and Figure 4(a), the difference between using drop priority or a combined approach becomes clearer. When in Figure 4(a), the fidelity stays higher even at the cost of loosing temporal resolution due to the defined minimums.

When looking at second 14 in the Figures the effect of the minimums becomes clearer. The result of the first approach (Figure 1) is a 4CIF slide show (1.875 FPS), while in the second approach (Figure 2) the user receives a video in CIF resolution, at high quality and with 15 frames per second. The algorithms with layer drop priority (Figure 3 and Figure 4) keep 30 FPS although at the cost of reduced quality, which corresponds to the user request. Moreover, with the third and the fourth approach less bandwidth is required compared to the first and the second approach, while perfectly matching the user preferences using the fourth approach. This shows that both layer drop priority and minimum values are relevant for an increased QoE at reasonable bandwidth utilization.

5. CONCLUSION AND FUTURE WORK

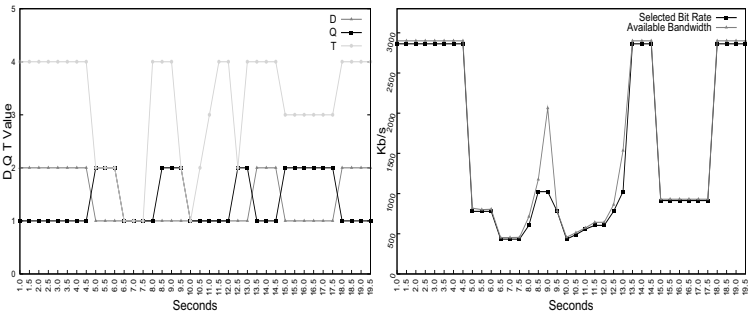
In this paper an approach for user-driven H.264/SVC video adaptation was described. After providing an introduction into the topic, we briefly presented state of the art and provided a motivation for our approach. Consequently, we presented our proposal and evaluated it. The evaluation showed the effectiveness of our approach and compared it to alternative approaches. It can be concluded that adaptation which uses our approach enables a higher QoE compared to existing adaptation mechanisms at the same bit rate.

Several items for future work can be seen. While



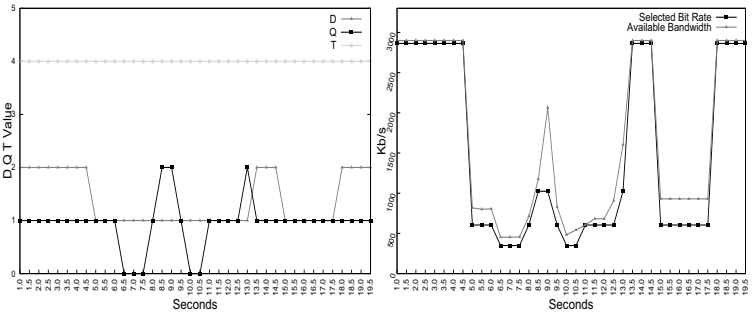
(a) DQT selection (b) Bit rate evolution

Figure 1: Best effort approach



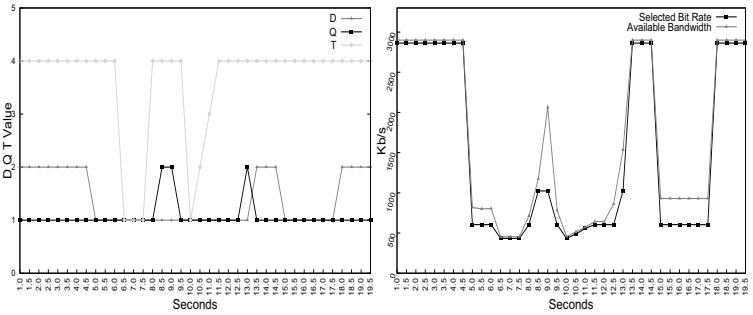
(a) DQT selection (b) Bit rate evolution

Figure 2: Minimal values approach



(a) DQT selection (b) Bit rate evolution

Figure 3: Layer drop priority approach



(a) DQT selection (b) Bit rate evolution

Figure 4: Combined approach

it is obvious that involving the user in the adaptation process results in a higher QoE, subjective tests need to be performed to be certain. An additional research item regarding the QoE is to research the user interface, i.e., how exactly can the user provide his preferences during the consumption of the video. A bad user interface could reduce the overall QoE of the user, thus negating the benefits of our approach. Finally we did not define how exactly to transmit the user preferences to an adaptation node, e.g., in an RTP streaming scenario.

6. ACKNOWLEDGEMENTS

This work is supported by the "Ministerio de Industria, Turismo y Comercio Español (MITYC)" and the "Österreichische Forschungsförderungsgesellschaft mbH (FFG)" in the context of the Celtic SCALNET (CP5-022) project.

7. REFERENCES

[1] T. Wiegand, G. Sullivan, H. Schwarz, and M. Wien, editors. *ISO/IEC 14496-10:2005/Amd3: Scalable Video Coding*. International Standardization Organization, 2007.

[2] T. Wiegand, J. Ohm, G. Sullivan, and A. Luthra. Special Issue on Scalable Video Coding - Standardization and Beyond. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(9), September 2007.

[3] A. Hutter, P. Amon, G. Panis, E. Delfosse, M. Ransburg, and H. Hellwagner. Automatic Adaptation of Streaming Multimedia Content in a Dynamic and Distributed Environment. In *Proc. ICIP*, Genova, Italy, September 2005.

[4] I. Kofler, R. Kuschnig, and H. Hellwagner. In-network Real-time Adaptation of Scalable Content on a WiFi Router. In *Proc. 6th CCNC*, January 2009.

[5] R. Kuschnig, I. Kofler, M. Ransburg, and H. Hellwagner. Design options and comparison of in-network H.264/SVC adaptation. *Journal of Visual Communication and Image Representation*, 19(8):529–542, December 2008.

[6] C. Abhayaratne, E. Izquierdo, M. Mrak, and S. Tubaro. Special Issue on Scalable Coded Media beyond Compression. *Signal Processing: Image Communication*, 24(6), July 2009.

[7] I. Kofler, C. Timmerer, H. Hellwagner, A. Hutter, and F. Sanahuja. Efficient MPEG-21-based Adaptation Decision-Taking for Scalable Multimedia Content. In *Proc. 14th Multimedia Computing and Networking Conference*, San Jose, USA, January 2007.

[8] A. Vetro and C. Timmerer. Digital Item Adaptation: Overview of Standardization and Research Activities. *IEEE Transactions on Multimedia*, 7(3):418–426, June 2005.

[9] ns-2 Network Simulator. Technical report, June 2009. http://nsnam.isi.edu/nsnam.

[10] M. Handley, S. Floyd, J. Padhye, and J. Widmer. TCP Friendly Rate Control (TFRC): Protocol Specification. Technical report, Internet Engineering Task Force, January 2003. Standard, RFC 3448.

Enabling distributed fault diagnosis in multicast-based IPTV delivery

Jorma Kilpi
VTT
P.O.Box 1000, FI-02044
Finland
Jorma.Kilpi@vtt.fi

ABSTRACT

We describe a monitoring approach which utilize QoE metrics and enables fast fault diagnosis in multicast based IPTV delivery.

1. INTRODUCTION

The IPTV quality is jointly affected by network dependent and application specific factors. The network dependent part include IP packet loss and jitter. A too large IP packet jitter translates into losses in the playback buffer [6]. The application specific factors include the video codec and coding bit rate, loss recovery technique, packetization scheme and content characteristics [6]. When the *Quality of Experience (QoE)* that an end user meets is insufficient there should be some relatively fast method to diagnose whether the reasons are in the network or in the application specific factors or, perhaps, in their joint incompatibility. In this paper we only focus on monitoring of multicast based delivery in an operational network.[1]

2. MOTIVATION

Figure 1 below describes a real-world example of IPTV delivery. There are four participants. Digita is a Finnish governmental company that takes care of TV distribution. The *IPTV Service Provider* is a commercial company which sells IPTV to the end users but buys the IPTV transport, the multicast delivery, from *the Transport Operator*. The fourth participant is a different operator who owns the copper cables over which the ADSL connections to the end users are made. The Transport Operator takes care of the IPTV multicast delivery. The IPTV Service Provider is a customer of the Transport Operator. An end-user is a customer of IPTV Service Provider.

We will adopt the point of view of the Transport Operator. The Transport Operator receives *Transport Streams (TS)* of all IPTV channels from the Headend, see Figure 1, and delivers each channel to all current subscribers of the channel. There are hundreds of channels available but not necessarily watched all of the time. Only those channels are delivered that are actually watched at the time. Between the Headend and DSLAMs the delivery is Ethernet multicast and between a DSLAM and the ADSL modem of a subscriber the delivery is over ADSL.

A problem identified is that it is not always easy to diagnose sufficiently fast whether the reason(s) for the bad QoE of customer(s) is (are)

1. Already before multicast transmission? This is not typical but seems to happen every now and then. It is always possible that the TS that the Transport Operator gets from the Headend is already damaged. It is also possible that the TS is formally correct but the encoded program that the TS contains is damaged. There is an indicating pattern that suggest whether the problem is already before the multicast transmission. Namely, many customers all over the city claim about bad QoE during a short period.
2. In the Ethernet part of the multicast tree? This seems to be quite seldom a problem since overdimensioning in this particular network should be sufficient even if all possible channels were delivered at the same time. Also, this is the part of the network that is solely under the control of the Transport Operator.
3. In the ADSL connections? This is very typical. The lengths of ADSL links vary and may be close to a practical maximum for the required bandwidth. Moreover, the interference with other customers may occur during the busy hours.
4. In the home network? This is also very typical. Home and, more generally, end customer network configurations vary a lot.

Especially, the question of interest in this study is whether the multicast delivery works correctly. Given the problem described above we suggest the following solution:

1. Measure *the availability of the TS* which carries a channel. (We will describe what we mean by the phrase 'availability of the TS' below.)
2. For a given channel, make the availability measurement continously at the root and then, periodically, multicast the measured availability value to all leafs (and nodes) of the multicast tree of the given channel's delivery.

[1] A longer version of this paper is available from `http://iplu.vtt.fi/multicast.pdf`

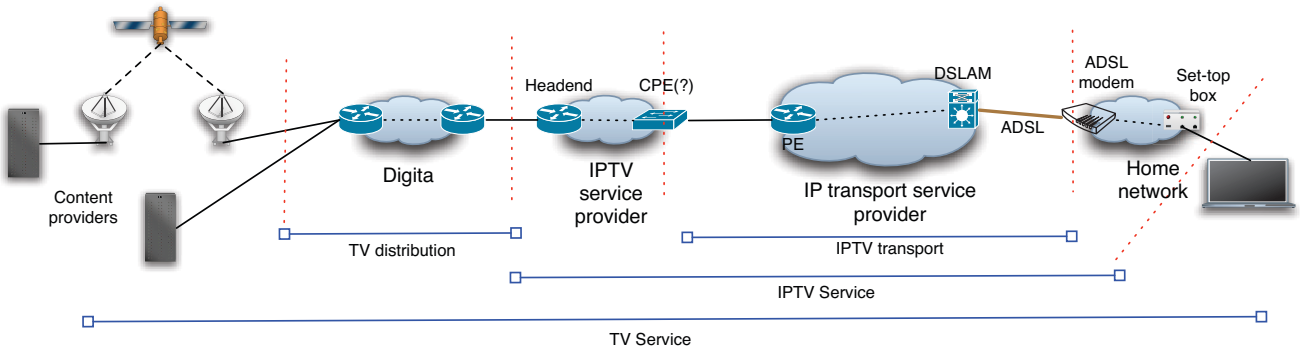


Figure 1: A real-world example of multicast based IPTV delivery.

3. At any node or leaf of the multicast tree that is of interest, make the same availability measurement of the TS, receive the corresponding availability measurement from the root and *compare* these two values.

3. BACKGROUND

MPEG-2 Transport Stream (TS) is introduced in [2]. TS is a coding syntax which is necessary and sufficient to synchronize the decoding and presentation of the video and audio information. The aim is also to ensure that data buffers in the decoders neither overflow nor underflow. In the case of a TV channel, voice, audio, subtitles and other ingredients each form a separate *packetized elementary stream (PES)* and these PESs are then multiplexed into a single TS.

The paper [5] discusses QoE assessment of IPTV. Three types of model are considered in [5]: *media-layer models*, *parametric packet layer models* and *bit stream layer models*. Media-layer models are essentially *Mean Opinion Score (MOS)* type of models where the knowledge of the human visual system is utilized in order to predict the subjective quality of video. Media-layer models can further be classified according to whether they use *full-reference (FR)*, *reduced reference (RR)* or *non-reference (NR)* methods. FR and RR methods utilize the original source video and its processed counterpart. NR methods operate solely on the information from the processed signal and, therefore, are most useful in an operational network [5].

The packet layer models are based on the packet header information only. They have the drawback that they do not look at the payload information. Thus, it is impossible for them to take into account the dependence of quality on audio and video content [5]. Furthermore, one can ask why to measure and process the IP-layer data, IP packet loss and jitter, since the inference from this information to the QoE is either trivial or extremely difficult.

From the Transport Operator point of view the most important thing is that the multicast delivery is successful. In order to be able to say whether the multicast delivery was successful one must be able to decode, that is, to process the TS at any leaf (node) of the multicast tree. This kind of reasoning suggests that the essential thing to be measured must be something that is able to take into account the requirement of continuous demultiplexing and decodability of the TS, taking QoE into account. This leads to bit stream layer model considerations.

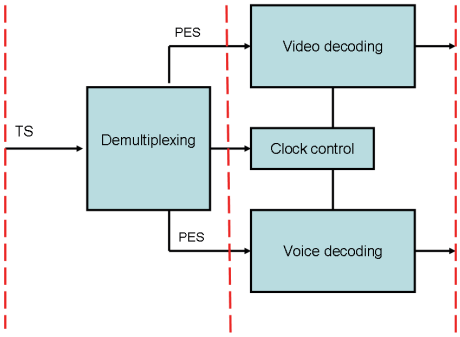


Figure 2: Decoding process of TS.

Measurement and analysis of the MPEG-2 TS is defined in [1]. The recommendation [1] also describes a number called *Unavailability Time (UAT)* which attempts to measure whether the TS is available or not. The definition of [1] is not applicable as such since it is not defined for IPTV. However, the main idea of the UAT number is worth of considering. We will discuss about what properties an optimal UAT number should have. These properties are motivated from what properties the UAT number of [1] has. Apparently, only minor changes or clarifications seem to be required.

Consider the decoding of TS as depicted in Figure 2. The availability control must, of course, be able to check that demultiplexing works correctly. This is based only on the information on the TS packet headers. In addition to this, the concept of TS availability should provide some level of guarantee that the decodings of the PES streams will work. Two-state state machines can be defined which measure the working/not working status of each of the processes (demultiplexing, video decoding, voice decoding, clock control) and transitions (TS, PES) in Figure 2. These include buffer under/overflow situations. The transitions between the working/not working states must be defined exactly, then they can be used to classify time periods into good and bad periods. Individual lower level state machines that measure the working/not working status of the demultiplexing and decoding processes of Figure 2 work in microsecond scale. It means that these state machines can be used to classify

milliseconds into good or bad milliseconds according to the condition that the state machine does not spend too many microseconds in a not working state. Then, with the help of these lower level state machines, a top level state machine can be defined which classifies seconds into good or bad seconds depending on whether or not all the lower level state machines are working sufficiently well or not in a millisecond scale.

For example, [1] speaks about Severly Errored Seconds (SES) and then the state transition from available state to unavailable state can be defined if there are more than M SESs in a window of size T seconds. The transition from unavailable state to available state occurs only if there are no SESs in a T-window. The M and T are parameters which must be tuned *according to user experience*. Once there is a clear method to classify each second as good or bad then the unavailability time could be defined in the same way as in [1]. Every second that is spent in the unavailable state increases the UAT number, in the available state the UAT number does not change.

The main point of view is that an optimal UAT number is a single number that is directly comparable and understandable without any further processing.

4. DETAILS OF THE SUGGESTED SOLUTION

We will now provide more details to the suggested solution. The main idea is that the comparison of the availability of TS, when measured at the root and at a leaf (or at any other node) of the multicast tree with sufficiently similar decoding parameters will automatically suggest whether the multicast delivery from the root to the leaf (node) is working correctly. For example, if the observed availabilities are sufficiently close to each other it implies that IP packet loss and jitter have not been too serious. The availability of TS cannot improve during the delivery, but the comparison measures how much the availability was degraded.

The availability of the TS must be *measured* at the root. One must not assume that it is automatically 100% (or that UAT is 0%) at the root. In order to get this measured value at any node of the multicast tree it must be delivered alongside of the tree. This can be a separate multicast tree that spans only those nodes that are willing to receive this information or it could be sent automatically to every member of the multicast group and those nodes which do not use it simply ignore it. The amount of information that is required to be sent is very small: we suggest periodic sending of one multicast packet per minute where the payload of the packet contains 2-5 numbers. These numbers would be UAT values of the last 1, 5 and, perhaps, last 60 minutes and some synchronization information.

The receiving node must somehow synchronize itself with the root so that the availability of the TS is measured during the same period. This synchronization could be based on the information in the stream itself, and transmitted together with the UAT values, or it could be based, for example, on NTP.

As a conclusion we argue that if the monitoring scheme described in this paper is implemented then we can answer the problem stated above. Namely, whether the reason for the bad QoE is

1. Already before multicast transmission? If the TS is formally correct but the content is damaged this could be partially solved by making a NR based MOS measurement at the root. (If the operator finds it worthwhile to try to react before the indicating pattern exists). If the TS itself is damaged then the TS availability measurement should notice it.
2. In the Ethernet part of the multicast tree? This is solved by comparing the TS availability measurement values between the root and the DSLAM.
3. In the ADSL connections? This is solved by comparing first measurement values between the root and the DSLAM and between the root and the ADSL modem. Then, comparing these comparisons allows to conclude about ADSL quality.
4. In the home network? The availability (and possible MOS) measurement at the root suggests whether the problem is already before the multicast transmission. Then, comparison of availability of the TS between the root and the ADSL modem allows to conclude about whether the problem is in the multicast delivery. The remaining possibility is that the problem is in the home network.

Pros of the suggested solution include

- The suggested scenario enables relatively fast fault diagnosis and fault localization. First results can be obtained in few minutes from the beginning of the measurement.
- The measurement process can be maintained as long as the TS-stream exist.
- Actually, the suggested solution could be applied into any type of streaming data transfer, not necessarily multicast nor IPTV.

Cons include

- It may not be applicable to more than only to a few simultaneous multicast trees at the time.
- It may not be applicable to pay charged channels that have an additional encryption.

We are seeking for research collaborators and companies to help us in prototyping the suggested monitoring scheme.

5. REFERENCES

[1] ETSI. *Digital Video Broadcasting (DVB); Measurement guidelines for DVB systems.*

[2] ISO. *Information technology–Generic coding of moving pictures and associated audio information: Systems.*

[3] I. Katzela, A. T. Bouloutas, and S. B. Calo. Centralized vs. distributed fault localization. In *Proceedings of the fourth international symposium on Integrated network management IV*, pages 250–261, London, UK, UK, 1995. Chapman & Hall, Ltd.

[4] M. Steinder and A. Sethi. A survey of fault localization techniques in computer networks. *Science of Computer Programming*, (53):165–194, 2004.

[5] A. Takahashi, D. Hands, and V. Barriac. Standardization Activities in the ITU for a QoE assessment of IPTV. *IEEE Communications Magazine*, February 2008.

[6] S. Tao, J. Apostolopoulos, and R. Guérin. Real-Time Monitoring of Video Quality in IP Networks. *IEEE/ACM Transactions on Networking,*, 16(5):1052–1065, October 2008.

Spatio-Temporal Quality of Experience Trade-offs for Mobile Imaging Applications

Muhammad Imran Iqbal
and
Hans-Jürgen Zepernick
Blekinge Institute of Technology
School of Engineering
372 25 Ronneby, Sweden
mii@bth.se

Markus Fiedler
and
Junaid Shaikh
Blekinge Institute of Technology
School of Computing
371 79 Karlskrona, Sweden
mfi@bth.se

ABSTRACT
In this paper, we focus on mobile images given in JPEG2000 format and evaluate their spatio-temporal Quality of Experience (QoE). In particular, the trade-off between spatial QoE supported by progressive source encoding and temporal QoE related to network delay as well as initial transmission delay is examined.

Categories and Subject Descriptors
H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*Performance measures*

General Terms
Quality of Experience (QoE)

Keywords
Spatial QoE, temporal QoE, spatio-temporal QoE

1. INTRODUCTION
The advent of advanced handheld devices such as smart phones resulted in an increased demand for mobile multimedia services including mobile TV, mobile video, mobile streaming, and mobile imaging. The quality of these services as perceived by the user can be measured in both spatial and temporal domain. The former relates to structural quality in terms of visual artifacts while the latter comprises of impairments caused by delays. In this paper, we focus on mobile images given in the JPEG2000 format and evaluate their spatio-temporal Quality of Experience (QoE). In particular, the trade-off between spatial QoE supported by progressive source encoding and temporal QoE related to network delay as well as initial transmission delay is examined.

The remainder of the paper is organized as follows. In Section 2, we provide a brief description of JPEG2000 features. The notion of QoE in terms of spatial, temporal and resultant spatio-temporal aspects is discussed in Section 3. Finally, numerical examples are given in Section 4

2. FUNDAMENTALS ON JPEG2000
The development of JPEG2000 aimed at providing increased coding performance and to alleviate drawbacks associated with JPEG such as blocking artifacts. JPEG2000 offers a number of advanced features including multiple resolution representation of an image and bitstream organization mechanisms to provide progressive decoding by quality and/or by spatial resolution [5]. In the context of this paper, we are primarily interest in the feature of progressive decoding by quality when receiving the bitstream over time.

The bitstream released by the arithmetic encoder represents the coded image data and is arranged into packets along with associated headers. The syntax of the resulting codestream comprises of headers, coded image data, and markers. Figure 1 illustrates the organization of the JPEG2000 codestream. It consists of a main header followed by tilestreams and termination by an end of codestream (EOC) marker. Similarly, each tilestream comprises of a tile header and a packstream. Eventually, each packet in the packstream consists of a packet header and carries portions of the actual image bitstream as packet data.

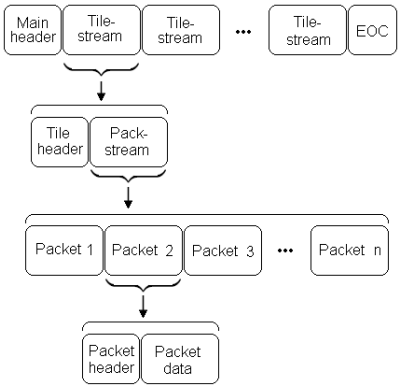


Figure 1: Organization of a JPEG2000 codestream.

3. QUALITY OF EXPERIENCE

Recently, the importance of taking care of user satisfaction with service provisioning has been realized. The related QoE concept covers end-to-end Quality of Service (QoS) parameters, factors such as coverage or support, and subjective factors such as expectations of and experiences with the service. In the sequel, we provide analytical expression that can be used to calculate the predicted mean opinion scores (MOS) of spatial and temporal quality. A simple multiplicative concatenation of these components is then used to deduce the overall QoE.

3.1 Spatial QoE

In view of the findings reported in [6], we take advantage of the fact that spatial QoE of a JPEG2000 image progressively improves with each packet received. In order to quantify this improvement, we utilize the well-known structural similarity (SSIM) index [8] and the following expression for its mapping to predicted MOS [1] [2]:

$$QoE_{\text{spatial}}(D) = 13.91 \exp\{1.715 \times \text{SSIM}(D)\} \qquad (1)$$

where D denotes delay and is used here to indicate the dependence on the progressively available packets over time.

3.2 Temporal QoE

Similar as with spatial artifacts, the temporal impairments inflicted by the network such as loss, jitter, reordering as well as impact of application and network stack can be modeled by a network utility function. Specifically, we can model the impact of network-caused delivery delays D of web content on user perception by an exponential function [7]:

$$QoE_{\text{temporal}}(D) = 0.97 \exp\{-0.15 \times (D/\text{s})\} \qquad (2)$$

This means that each extra second of waiting time "costs" approximately 15 % of the user ranking.

3.3 Spatio-temporal QoE

In [3], [4], we have quantified the impact of network-level disturbances on user perception through a network utility function U_{net}, which interrelates user-perceived utility U_{in} of a service without using a network and user-perceived utility U_{out} of a service after using a network through

$$U_{\text{out}} = U_{\text{Netw}} \cdot U_{\text{in}} \qquad (3)$$

Applying (3) to our case of progressive image download, we can identify the spatial quality that improves progressively using (1) and (2) to express the overall QoE as

$$QoE_{\text{spatio-temporal}}(D) = QoE_{\text{spatial}}(D)\, QoE_{\text{temporal}}(D) \qquad (4)$$

4. NUMERICAL EXAMPLES

In the sequel, we present numerical examples in order to illustrate system design trade-offs between spatial QoE and temporal QoE. The following parameters have been chosen for producing the results shown in Figs. 2 and 3:

TCP Header Length:	28 bytes
TCP Payload Length:	100 bytes
TCP Packet Length:	128 bytes
Data Rate:	64 kbps and 384 kbps

Figure 2 shows different aspects of the spatio-temporal QoE for a data rate of 64 kbps and image sample 'Lena'. As can be seen from Fig. 2(b) the spatial QoE increases with the delay as the number of packets available to reconstruct the image increases progressively. On the other hand, the temporal QoE decreases exponentially with longer delay as this translates to waiting time for the end-user. Accordingly, the resulting spatio-temporal QoE as the product of the component QoEs reveals an optimal delay value that maximizes the resulting QoE and supports an optimal trade-off between spatial and temporal QoE. Figure 2(c) presents the resulting QoE for different initial delays and illustrates the decrease of QoE due to this additional waiting time. It should also be noted that during this time no packets are transmitted and hence no information is transferred that could be used for image reconstruction. In view of the above results, Fig. 2(d) shows the maximum spatio-temporal QoE for optimal delay of 2.21 s but different initial delays.

Figure 3 provides results for spatio-temporal QoE using an increased data rate of 384 kbps and image sample 'Barbara'. Similar behavior as for the scenario with the lower data rate can be observed but the resultant QoE reaches a significantly larger maximum value. This can be attribute to the fact that the complete image is progressively transmitted during the respective delay while the actual delay is rather short due to the high data rate. In other words, only little degradation of temporal QoE is incurred while the spatial QoE reaches the maximum possible level at optimal delay of 0.74 s.

5. REFERENCES

[1] U. Engelke. *Perceptual Quality Metric Design for Wireless Image and Video*. Licentiate thesis, Blekinge Institute of Technology, Sweden, 2008.

[2] U. Engelke, M. Kusuma, H.-J. Zepernick, and M. Caldera. Reduced-reference metric design for objective perceptual quality assessment in wireless imaging. *Signal Processing: Image Communication*, 24(7):525–547, Aug. 2009.

[3] M. Fiedler et al. The network utility function: A practical concept for assessing network impact on distributed services. In *International Teletraffic Congress*, pages 1465–1474. Beijing, China, Sept. 2005.

[4] T. Hoßfeld, A. Binzenhöfer, M. Fiedler, and K. Tutschku. Measurement and analysis of Skype VoIP traffic in 3G UMTS systems. In *IPS-MoMe*, pages 52–61. Salzburg, Austria, Feb. 2006.

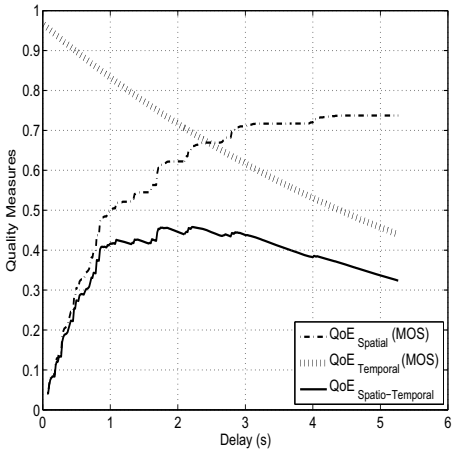
[5] International Organization for Standardization. JPEG2000 image coding system: Wireless. *ISO/IEC 15444-1:2004(E)*.

[6] M. I. Iqbal, H.-J. Zepernick, and U. Engelke. Error sensitivity analysis for wireless JPEG2000 using perceptual quality metrics. In *International Conference on Signal Processing and Communication Systems*, pages 1–9. Gold Coast, Australia, Dec. 2008.

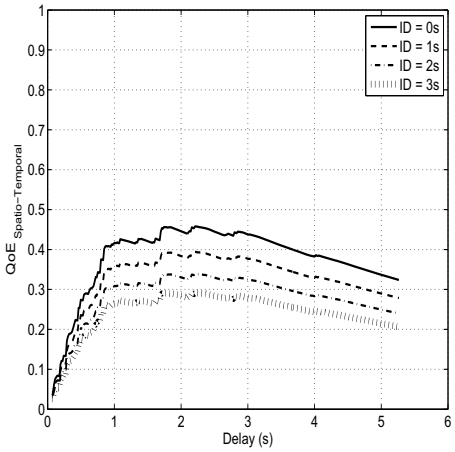
[7] J. Shaikh, M. Fiedler, and D. Collange. Quality of experience from user and network perspectives. *Annals of Telecommunications*, 65(1-2):47–57, Jan./Feb. 2010.

[8] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. on Image Process.*, 13(4):600–612, April 2004.

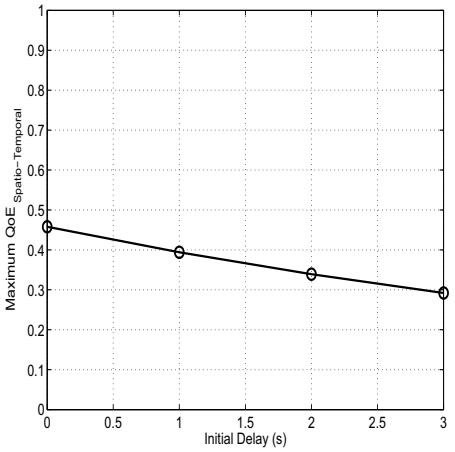
(a)



(b)



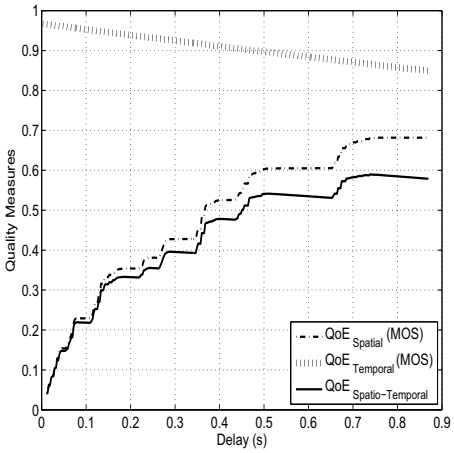
(c)



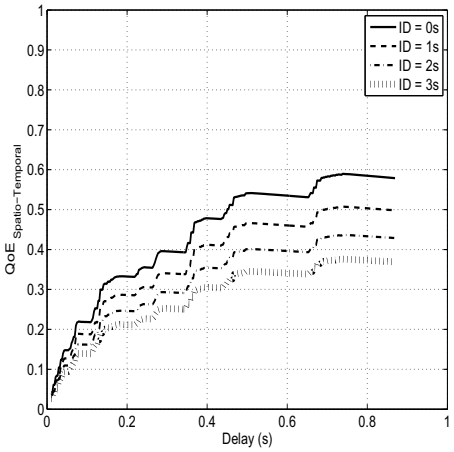
(d)

Figure 2: Spatio-temporal QoE for 64 kbps: (a) Image sample 'Lena', (b) QoE measures, (c) Resultant QoE, (d) Optimum QoE.

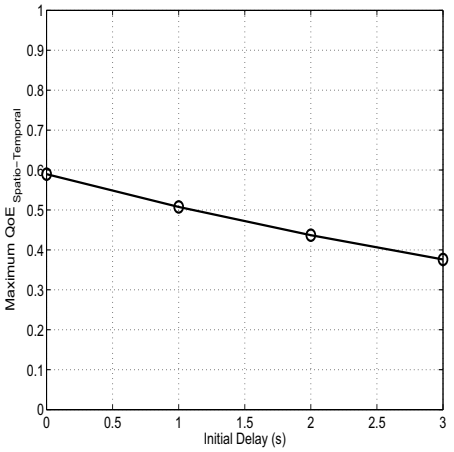
(a)



(b)



(c)



(d)

Figure 3: Spatio-temporal QoE for 384 kbps: (a) Image sample 'Barbara', (b) QoE measures, (c) Resultant QoE, (d) Optimum QoE.

WORKSHOP 4

Future Television: Integrating the Social and Semantic Web

Editors:

Dr. Lyndon Nixon (STI International, Austria)
Dr. Lora Aroyo (VU University Amsterdam, Netherlands)
Dan Brickley (VU University Amsterdam, Netherlands)
Libby Miller (BBC, UK)

Linking TV And Web Using Semantics, A NoTube Application

Balthasar Schopman
Vrije Universiteit
De Boelelaan 1081A
1081 HV Amsterdam, The Netherlands
schopman@cs.vu.nl

Davide Palmisano
Fondazione Bruno Kessler
Via Santa Croce 77
I-38122 Trento
davide@asemantics.com

Ronald Siebes
Vrije Universiteit
De Boelelaan 1081A
1081 HV Amsterdam, The Netherlands
rm.siebes@few.vu.nl

Chris van Aart
Vrije Universiteit
De Boelelaan 1081A
1081 HV Amsterdam, The Netherlands
chris@vanaart.com

Véronique Malaisé
Vrije Universiteit
De Boelelaan 1081A
1081 HV Amsterdam, The Netherlands
vmalaise@few.vu.nl

Michele Minno
Pro-netics
Viale Luigi Schiavonetti, 270
00173 Rome, Italy
michele.minno@asemantics.com

ABSTRACT

In contrast to TV, the Web is a highly dynamic environment that encourages active participation of users. Many Web services contain bits and pieces of data about users, such as on social networks, where people discuss what they like and dislike. The NoTube project aims to combine existing services and create a system that exploits user data, TV programmes and background data to produce accurate content-based recommendations.

Keywords

personalisation, integration TV and Web, Linked Data cloud

1. INTRODUCTION

With the hundreds of channels available via modern TV providers, content selection and dealing with the vast amount of TV-related information become significant challenges for users. TV metadata is created and distributed by a small group of people, as a result of the closed-source information exchange protocols that are the standard for providing *electronic programme guide* (EPG) data to users. The NoTube project puts the TV user back in the driver's seat by generating user profiles from data the user creates on the Social Web, such as tweets on Twitter,[1] and a network of friends on Facebook,[2] and in this way facilitating a personalized TV experience without an intrusive user profiling process.

As much as the Web encourages users to actively participate, TV remains a largely passive experience. NoTube envisions a common ground for the Web and TV, where people actively contribute while watching TV, by discussing and recommending programmes. NoTube aims to make TV a community-enhanced platform by overcoming the barriers of the physical location of the TV set, which makes TV a single-user (or single-family) experience.

People often have several clusters of personal data on the Web, such as their profiles on social networks, or ratings of videos on YouTube[3] and IMDB.[4] Analogously, there are many isolated clusters of broadcast data on the Web, such as broadcast data on EPGs and background information on Wikipedia. Within the NoTube vision context, we speculate that the conjunction of all these bits and pieces of data would provide accurate information of someone's interests and be suitable for generating relevant recommendations on TV broadcasts. Further, we assume that the Semantic Web technologies provide important building blocks for realizing this vision, as they enable the global identification mechanism of URIs and the means to define relations between data anywhere on the Web.

Due to its dynamic nature, the Web has been shaped to the desires of its users. TV technology is not as dynamic as the Web and has kept its linear programming over the last decade, partially due to technological constrains of the medium and its business model where advertisements remain the most important source of income for commercial TV stations. The strong interaction between the desires of users and technology has had its impact on the Web and as the gap between the Web and TV experience grows, we aim to translate features of the Web to TV, such as the personalised and community-based aspects.

In this paper we address one of the three applications that are developed in the NoTube project: 'TV in the 21st century'. Jana, our user, is socially active on the Web and does

[1] http://twitter.com/
[2] http://facebook.com/
[3] http://youtube.com/
[4] http://imdb.com/

not see the need to explicitly define her preferences or wait until she has used the recommender system long enough for it to learn her preferences. In the first use-case we deal with Jana's recommendations based on her online social activity. In the second use-case, Jana is interested in a programme and uses the 'I would like to know more'-button. Jana then gets information about this program, which contains links to Wikipedia, IMDB or other sources entries. Next this this, she also gets recommendations of related programmes. With the 'why'-button option Jana can see why each program has been recommended to her. As we consider enriched TV program descriptions, the reasons for recommendations often are based on interesting semantic relations between entities.

In this paper we describe how the NoTube project aims to build a recommendation system that uses existing Web services and shared background knowledge to collect, enrich and recommend TV data. In Section 2 we give an overview of the NoTube system. We describe the data sources and how we enrich them with semantic data in Section 3. Section 4 is about the user profiling aspect, which is crucial for the recommendation phase presented in Section 5.

2. NOTUBE SYSTEM OVERVIEW

The purpose of the NoTube system is to re-use user and TV data scattered around the Web to provide a personalised TV watching experience. With the huge amount of metadata on the Web about television programmes, movies, music, etc, and personal data users generate, e.g. on social networks and by rating movies on, for example, IMDB, it is possible to personalise a user's TV experience without having the user explicitly telling the system what her preferences are.

Figure 1 shows an overview of the NoTube System. The user registers for the NoTube system by providing information, such as her Facebook, Twitter and/or YouTube account, using the Web browser to access the NoTube web interface. We discuss four main clusters of services here:

- The *Beancounter*, a group of services that collects user data and generates a user profile from it, harvests and uses the user data from online user data repositories, which are used to generate a user profile. The Beancounter is addressed in detail in Section 4;
- The *Home Ambient*, a group of services that replace the old fashioned TV and antenna. For basic TV interaction, the user can either use a smartphone or a regular laptop as a remote control. The remote control communicates with the set top box providing the connection to the TV screen. Additional to basic TV interaction, the remote control can be used to rate, bookmark or find out more about each TV programme. The interaction between the user and the remote generates a user activity stream, which is processed by the Beancounter and stored for use further in recommendations;
- The *Data Warehouse* is a group of services that collect and enrich EPG data. Existing EPG harvesting services, such as XMLTV,[5] are used to obtain EPG data. The descriptions of TV programmes are then enriched by the Named entity recognition service, which identifies entities from the Linked Data cloud. The NoTube vocabulary alignment service identifies links between concepts of different vocabularies. The data warehouse services are described in detail in Section 3;
- The *Recommender* service is the nexus of the NoTube System - it generates recommendations by combining the output of the data warehouse and the Beancounter in the Home Ambient context. Recommender strategies that use Semantic Web technology and the Linked Data cloud to reason with background data of TV programmes and user interests, are discussed in Section 5.

[5] http://xmltv.org/

3. DATA SOURCES, SCHEMAS, VOCABULARIES AND ALIGNMENTS

In this section we present the datasources, schemas, vocabularies and alignments that we identified as relevant for the enrichment of TV content. By disclosing data of heterogeneous sources linked to existing RDF schemas and vocabularies we hope to offer a uniform, standardised and semantically formalised set of data that can be integrated in the Linked Data cloud and can be processed by Semantic Web tools.

3.1 Data sources

We transform various existing (TV-related) data into RDF and map it to existing RDF schemas for describing content relevant to the TV domain. The original data is often scraped from webpages and transformed to semantic data. We adopt the TV-Anytime schemas and vocabularies as the standard representation of EPG data, since TV-Anytime provides a well-structured and versatile format.

Another source of data are the user profiles and user activities. For example, a user profile on twitter or Facebook may contain information about the programmes being watched, liked, not liked. User interests could be categorised in genres (e.g. Science-fiction). By collecting the genre hierarchies from different sources (e.g. IMDB, YouTube, tvgids.nl[6]) and converting them into RDF concepts and defining SKOS relations between them, we are able to connect instance data that comes from these different sources.

The choice of data sources is driven by some typical use-cases (provided by project partners, e.g. BBC) which are envisioned to be realistic within the near future. We link EPG data from broadcasters, social network data about user interests, behaviours and contexts, user TV watching activities, and sources that provide relevant background information about TV material. Below a short overview of the sources that we incorporate in our transformation process.

3.1.1 EPG data

An EPG is composed of two parts: content descriptions and broadcast description. Content descriptions contain static data about television programmes such as a brand name, e.g. EastEnders description or plot summary, type of programme, e.g. series, movie or news, genre(s), e.g. drama, actors, directors, recording data, etc. Broadcast description

[6] http://tvgids.nl/

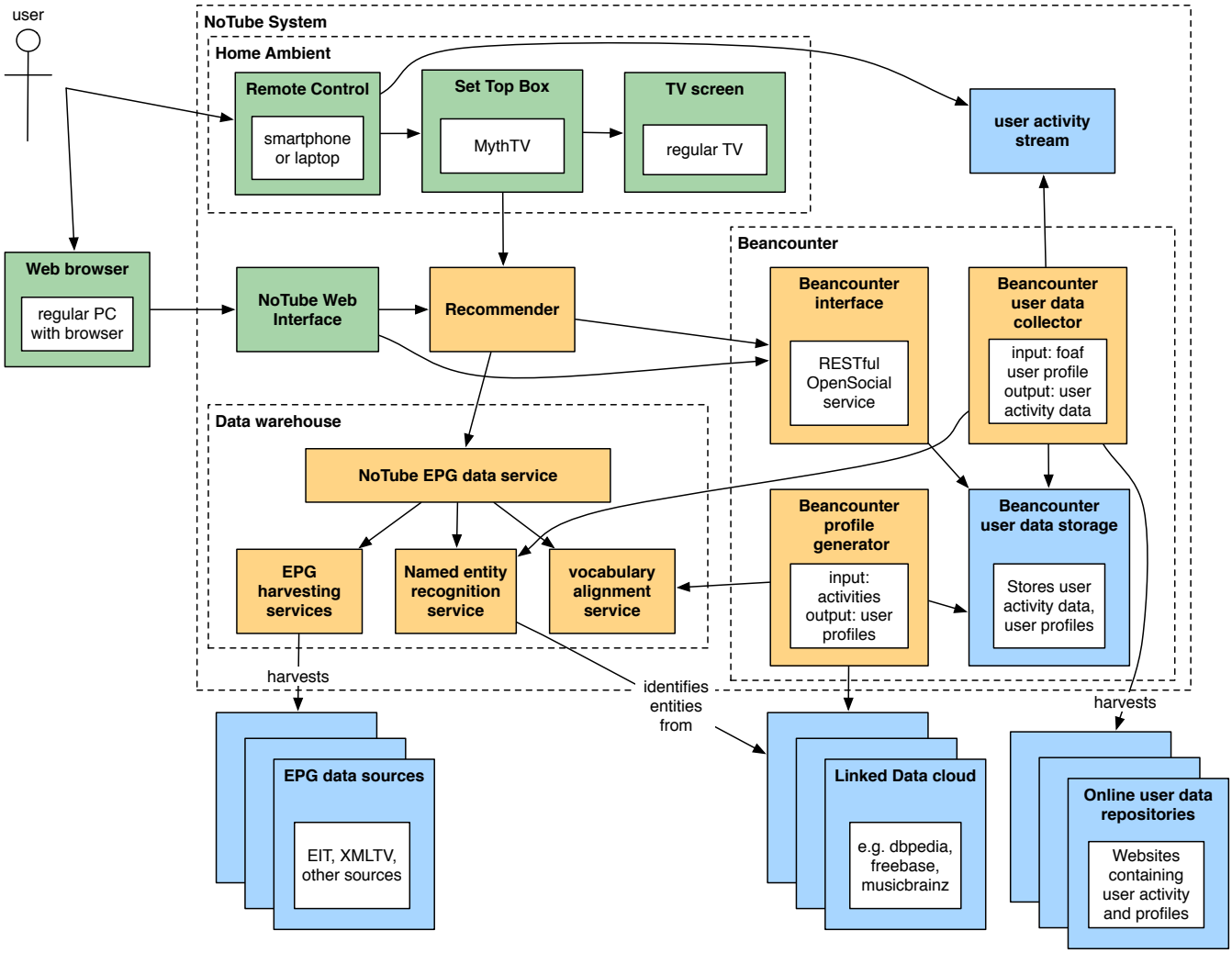


Figure 1: Overview of the NoTube System. Boxes represent components providing services; links represent service calls, unless specified otherwise. Green services are services that users have direct interaction with, blue services are interfaces to data repositories. Orange services provide algorithms to enhance the TV experience.

is expressed by variable data, such as channel (e.g. BBC ONE), format (e.g. 16:9) and broadcast media (e.g. digital television).

In general, EPG data is produced by broadcast companies. Some broadcasts companies, such as the BBC, have made their EPG data machine-readable and publicly available.Other EPG data is harvested from websites using existing tools, such as XMLTV, and converted to a machine-readable format.

3.1.2 *Vocabularies and schemas*

The data sources described above are often already annotated with a fixed set of concepts. For example, EPG data from the BBC[7] is annotated with BBC-defined genre hierarchy and IMDB categorised TV-series and Films into a similar set of genres. Vocabularies can be domain-independent, for example Princeton WordNet[8] provides a set of lexical concepts that match words (in e.g. English) and provides semantic relations between those words. Such vocabularies can be used to annotate domain-specific data. For example, the description of a movie could be a set of WordNet concepts. For some datasets the Semantic Web community already converted the vocabularies and schemas in RDF, like the BBC Programmes ontology,[9] the TV-Anytime schemas and vocabularies,[10] W3C WordNet.[11] To cover multiple perspectives, we also created (extracted) additional genre vocabularies for sources like YouTube and defined vocabularies for TV related user activities, e.g. "watching, reviewing, rating".

[7] http://bbc.co.uk/programmes/

[8] http://wordnet.princeton.edu/

[9] http://www.bbc.co.uk/ontologies/programmes/

[10] http://http://www.etsi.org/website/technologies/tvanytime.aspx

[11] http://http://www.w3.org/TR/wordnet-rdf/

The choice of vocabularies and schemas automatically follows from the choice of the aforementioned datasets in combination with the requirements from the enrichment strategies and recommender algorithms. Enrichments of EPG metadata are used to provide the end users with extra information about the content in which they are interested. For example, a scheduled broadcast of a movie could have an enrichment that enumerates the main actors together with the pointers to IMDB. Recommender algorithms that fall in the category of content-based filtering algorithms use content descriptions of the items for determining the relevance to the users. To give a simple example, when a user often watches content annotated with the Western concept, then other content annotated with the same or related concepts may be interesting her.

Below a short list of some of the RDF schemas and vocabularies that are used within the project:

- **TV-Anytime** consists of a wide range of templates to describe TV related metadata. For example EPG data, types of broadcasts and distributions, data stream formats, broadcaster information etc. Given its high quality, focus on standardisation and richness, we choose TV-Anytime as the default schema (and vocabulary) to annotate the EPG data in NoTube.
- **BBC-Programme ontology** is developed to describe the BBC EPG information, but can also be used for EPG data from other sources. One of the datasources that is used within NoTube is EPG data from the BBC which is already annotated with this schema. In the next section we describe how we align TV-Anytime with the BBC-Programme ontology.
- **Linked Data cloud** contains a wide variety of related and less related schemas and vocabularies. Most relevant are those that formalise social network information, content descriptions and user activities.
- **Web-services** and brokering of these services play an important role in NoTube. These building blocks are combined to provide the functionality that is required by the use-cases, but also can be used by developers outside the NoTube consortium. Especially for the latter requirement, using Web service schemas like WSMO-lite provide a standardised means to describe the functionality of these building blocks.
- **Event model and Activity model**. There are several types of typical activities in the TV-domain. A user views and comments on video content. A broadcaster shows content at certain times and in specific formats. The content of the videos often contains activities. For example, footage of the US elections may contain activity of Obama giving a speech somewhere. An important task within NoTube is to specify these typical activities and create the appropriate schemas and vocabularies. Reusing existing vocabularies to describe parts of the activities is our first choice.
- **SKOS** is a schema to represent vocabularies and relations between entities of those vocabularies in a domain independent way. For example, we use SKOS to express the relations between genres that are defined in different vocabularies.
- **FOAF** is a schema and vocabulary to describe personal information.
- **Genre vocabularies** are sets of genres combined with the relations, originating from the datasources used within the NoTube project. For example, the BBC-genres, IMDB-genres, Uitzending gemist, RAI, YouTube and FreeBase.
- **WordNet** is a large lexical database of English, developed at the Princeton University. English words are organised in sets of synonyms (synsets); synsets are organised in conceptual relationships (e.g. hierarchical relationships) and in lexical ones (e.g. antinomy).
- **UWN**[12] is a multilingual lexical database automatically constructed and linked to the Princeton WordNet. It contains over a million words for more than 200 languages; the coverage is, however, not consistent for all the languages. We have a dump dating from sept 2009 for the following languages: French, Italian, Turkish and Korean.

3.2 Enrichment

In the NoTube project we enrich broadcast data using the Linked Data cloud, by linking existing data source, e.g. DBPedia[13] (subject data), Yago (data about people), IMDB (data about movies) to broadcast metadata

With the metadata enrichment we enable finding interesting relations between programs, and subsequently between user's interests, as described in Section 5. Next to the metadata enrichment we also perform vocabulary enrichment by aligning vocabulary hierarchies, e.g. genres.

The enrichment process undergoes the following phases:

Phase A. entity recognition: based on a title from the broadcast and its metadata facets (e.g. genres, people involved, producers, actors, directors, writers and composers) we query existing Linked Data cloud for concepts, and then link broadcast metadata values to the concept URIs found in the Linked Data cloud vocabularies . In case of multiple entities with the same title (e.g. Matrix I, Matrix II and Matrix III) we use additional criteria, e.g. recording data in order to disambiguate the TV content.

Phase B. metadata enrichment: based on the found entities and their concept URI counterparts we:

- select relevant repositories, e.g. YAGO, DBPEDIA, IMDB, WIKIPEDIA, subtitle and various gossip sites,
- based on found entities extract relevant - textual - features, such as description, plot, scenario, subtitles, comments, critics

[12] http://www.mpi-inf.mpg.de/yago-naga/uwn/
[13] http://dbpedia.org/

Phase C. vocabulary alignment: we align content from different sources, e.g. genres into general format: TV-Anytime. Genre alignments are created to express the relation between the different genre vocabularies. These alignments are either created manually or automatically.

4. USER PROFILING USING SOCIAL ACTIVITY DATA

User modelling is probably one of the most investigated fields in the recent years. Many steps have been made since the first engineered approaches [5] came out. As a major key to lower the information overhead, various user profiling techniques have been employed to improve searches on the Web and to make the access to information personalised. In NoTube the user profiling component also plays a main role for the enabling a personalised access to TV contents.

Traditionally, consolidated user modelling representations fall under these four main typologies:

- Stereotypes - the user characteristics are modelled as a set of rules and facts (often following First Order Logic).
- Keyword-based - each keyword can be directly provided by the user or automatically extracted from the content she consumes. Keywords represent a topic of interest for the user. Basic construction methods are, mainly, derived from the classical information retrieval theory [1, 9]
- Semantic Network-based - a semantic network (graph), where the nodes represent keywords extracted from a document corpus and the weighted arcs represent their co-occurrences [6].
- Concept-based - the main difference with the Semantic Network-based profiles is that in a concept profile nodes are concepts and arcs are conceptual relationships between concepts.

Even if this classification is still valid and it covers all the different aspects and features of a generic user profiling system, attempting to use it to categorise the approach we followed in NoTube could be misleading. In this sense, it is important to remark that the approach adopted in the NoTube project about user profiling is grounded on a set of considerations that come out from the characteristics of the environment where it acts. More specifically, approach and algorithms have been heavily influenced by the following considerations:

- The huge amount of interactions that a user perform on the Web represents a powerful and extraordinary source for mining his interests and preferences. Even if several attempts already tried to infer a profile of a user starting from his behaviour on the Web [2, 3, 4], the NoTube approach focuses on the opportunities unleashed by the so-called Social and Real-time Web - where users discover, consume and share contents within their social graph, in a real-time manner and often using mobile devices. In such environments each user produces a rich and machine-readable flow of activities from which we can extract implicit and explicit information regarding his preferences.
- The access to these information flows is hindered by a plethora of different issues. The first phenomenon, often referred to as the Balkanization, refers to the fact that a user leverages different Web applications to satisfy her information needs. These applications employ different technology to expose their data, with different data formats, and different domains of applications. A general purpose social network Web application has different finalities from a social bookmarking tool or from a real-time microblogging service. These application domains naturally lead to different data models and different technology to access data but, most of all, they all could implicitly bring relevant information about user interests.
- Web resources implicitly carry information about things: a book, a movie, another Web resource or a person. Even if this consideration could lead us to identify a huge class of problems that are out of scope here, it is important to underline that once we declared that we are interested in mining the information on a specific user starting from the Web resources she consumes, then we must foresee an heuristic mechanism that could allows us, when possible, to identify Web resources referring to the same topics. it was already there in the traditional Web.
- In NoTube, we use Semantic Web technology to represent TV metadata. As a consequence, adopting RDF as a data model allows us to use vocabularies and ontologies for a powerful data integration. Secondly, it allows for disambiguation of topics, genres and various equivalent/similar concepts. Representing user profiles as RDF descriptions containing links to some shared concepts allows us to make content recommendations and discovery by following links and description contained in such ontologies.

User profiling systems can heavily differ in terms of technology and methodology they use to achieve the three main constitutive phases. Informally:

1. Collecting: the raw data about the user are collected. This phase strongly influences next phases in terms of what kind of user information could be represented.
2. Profiling: the gathered raw data are represented as instances of a suitable user model. The whole user model life-cycle is addressed within this stage: building, maintaining, updating the user model, as well as all that is needed to keep the profile quality and consistence, are performed within this step.
3. Exporting: the generated user profile is made suitable for an external application. The application will make use of it in order to produce personalised content delivery.

In the next two subsections we will describe how the considerations above impacted and influenced both algorithmic

and architectural aspects of our user profiling system. In the first section we describe the methodology adopted to collect the user raw data, while in the second one we will see how we build and maintain a coherent and updated user profile.

4.1 A Linked Data approach to the user modelling

We believe that the approach described here is novel: it uses both Semantic Web technology and Linked Data ontologies to build a user model that could solve issues raised by the previous methods. Some user modelling systems already leverage Semantic Web technology in order to perform personalisation activities. Interesting work done in a adaptive web-based system context [4] aims to represent the different dimensions of the user (i.e. her features, actions, context) as different planes of a multidimensional matrix and formalised by means of ontologies. This UMS represents the semantic dimensions of the matrix in OWL and uses SWRL44 as a rule language to infer a user profile. This approach is similar to the one adopted within NoTube, where, ideally, the matrix is replaced by a the Linked Data cloud.[14]

4.2 Collecting and aggregating user data from the Social Web

Here we show how the user activity data is transformed by various mechanisms to achieve a twine of interlinked machine-readable data that is used to feed the user profiling algorithm. Our scenario considers a generic user who holds at least two accounts on different social web applications: Last.fm[15] and Glue.com.[16] Last.fm tracks the user to listening, sharing and live events activities. Glue.com acts like a user log, realized as a browser plugin. Glue.com makes available through a set of Web APIs an exhaustive set of Web resources the user visited, shared or liked, enriching them with an internal categorisation. The following sections show how we aggregate data from these sources, link the information to several Linked Data clouds and reason using the data in order to make explicit the user's interests in a user profile. The information aggregation is achieved through the identity resolution made against different ontologies. It will be described underlining its importance for the final user profiling objective. The following example shows how we can infer a set of interests, uniquely identified by Linked Data identifiers, starting from the tracks and the Web resources accessed by a user across two different social networks.

4.2.1 Aggregated activities in RDF from social networks

The Last.fm API response is a list of tracks the user recently listened to, providing for each of them the artist and the album UUIDs,[17] a name and other data that is out of scope here. The Glue.com response is a list of interactions that the user had with Web resources, typically html pages or anything that can be uniquely identified with an URL. Such interactions are described in term of the HTTP link, the title of the resource, the action the user performed on it (e.g. liked, shared, commented, bookmarked) and among others the category of the resource according to Glue.com internal hierarchy of concepts. Here we observe:

- Representation Heterogeneity: two different XML schemas are employed. Thus, we represent such data in RDF according to the Atom Activity Streams Vocabulary.
- Data Model Mismatch: next to the differences in syntax, there is also different kinds of implicit information the two services are bringing within their own data models. The main challenge here is trying to leverage such flat descriptions and integrate them on a semantic level. We can assume that the tracks showed by Last.fm are from artists in which the user has some interest and such interests also reflect on the way she browses on the Web with Glue.com, then it is possible to leverage such implicit information, making them part of her profile.

To uniformly represent user activity data of different sources in a single graph, the *ATOM Activity Streams in RDF vocabulary*[18] is used to represent user activities. To determine the objects of activity, we use a named entity recognition service similar to Phase A in Section 3.2. An alignment service is used to link the objects of different vocabularies, e.g. last.fm artists are linked to DBpedia entities and the BBC Music[19] catalog.

4.2.2 A SKOS-based user profiling algorithm

The data generated by the data collection and enrichment process, as described above, forms a potentially huge amount of activities for each individual user. The challenge is to derive general user interests from this set of user activities. We do so by using the *DBpedia SKOS vocabulary*, which is the semantic counterpart of the *Wikipedia Categories*.[20] The main idea is to use the skos:subject property values as identifiers for the user interests. If a user listens to bands or musicians sharing the same skos:subject, then it could be reasonable to infer that such subject represents an interest for that user. Moreover, the rich and complex SKOS hierarchy of DBpedia allows to extract a lot of other interesting information. For example, if a user is particularly interested in movies where a particular actor or actress played, more information will be available about this in the system, since it is highly probable that DBpedia contains some SKOS subjects describing this. Similarly, if a user listens bands originating from a specific geographical region then this could be represented with a skos:subject property could be a statement useful to perform recommendations of other bands and artists.

The developed algorithm has two aspects:

1. *The linkage of the resources across the Web of Data.* This is achieved by fetching the various SKOS subjects of the resources to be categorised. A recursive algorithm ranks a set of SKOS subjects according the

[14] http://linkeddata.org/
[15] http://last.fm
[16] http://getglue.com
[17] http://en.wikipedia.org/wiki/UUID
[18] http://xmlns.notu.be/aair/
[19] http://www.bbc.co.uk/music
[20] http://en.wikipedia.org/wiki/Help:Category

degree of representativity that they have on a specific resource URI given as input. Essentially, the degree of representativity of a SKOS subject is calculated comparing the SKOS subjects of its most linked resource with the one provided as input. For a complete and formal description, see [7].

2. Then such set of skos:subject is collected and counted, revealing the statistical side of the algorithm. Only subjects that have been counted more than a certain threshold value are considered as meaningful for the user profile.

The resulting user profile is a set of Weighted Interests.[21]

5. RECOMMENDATION STRATEGIES USING SEMANTICS

Using semantics in recommendation strategies has several advantages: it enables the use of background data, recommending items based on specific relations between properties and facilitates specifying a reason for recommendations in natural language.

By enriching the EPG data, as described in Section 3.1, links to semantic entities in the Linked Data cloud are added to the metadata of TV programmes. The interconnected entities in the Linked Data cloud allow for finding interesting relationships between entities, e.g. that two movies have been made by people that have a common interest in *film noir*. The relationship between entities are often typed, e.g. by SKOS[22] relationships. Since not all relationships between entities are considered interesting, e.g. the relation rdf:type Person is generally unremarkable, the types of the relationships must be taken into account during the recommendation process [10]. This can be done by using relationships of specific patterns and/or assigning a certain weight to specific relations.

Another advantage of using semantic data and typed relations is that the reason of a recommendation can easily be displayed in natural language. Both semantic entities and relationships often have a preferred label and optionally alternate labels assigned, which facilitate the conversion of a semantic relation to natural language, as shown by the Music Bore.[23] We expect that providing a reason for recommendations improves the trust of the user, which is an important factor in the user adaptation of recommendation systems [8]. For example, when a user does not understand why a certain documentary is recommended by looking at the title and description, but the recommendation system displays the link between this documentary and the user's interests, the user may be convinced of the significance of the recommendation and keep high expectations of future recommendations.

Recommendations can be made for specific user considering her user profile or just considering a selected content item. An example for the first case is: Jana has watched many movies on Boxee,[24] a media centre that allows users to automatically propagate their consumption and ratings of audiovisual content on social networks. She registered her Twitter account on Boxee, so a large amount of her user activity data is on Twitter. She uses for the NoTube service, because she does not see the need to tell the recommender system what she likes or wait for it to collect a sufficient amount of user activity data, since it is already on her Twitter account. When she turns on her TV, a recommendation of the movie 'A Serious Man' is shown, because it is broadcasted on BBC Two and matches her interest in 'black comedy' and her preference for films directed by the Coen brothers, which were both in her user profile that the Beancounter generated using the user activity data collected from Twitter.

An example for the second case is: When Jana is watching an episode of 'True Blood' this makes her curious about the series, so she picks up her smartphone to find out more about it using the NoTube application. When she presses the 'I want to know more' button the Wikipedia page is shown as well as some recommendations. One of the recommendations is the pilot of the series 'Six Feet Under', which she already knows. She is curious about the reason of the recommendation, so she presses the 'why?' button next to it and sees that: both series were created by 'Alan Ball' and they share two genres: 'black comedy' and 'drama'. She is happy to learn that the two series were created by the same man and continues by looking up information about Alan Ball.

Several open issues of recommender strategies that use semantics exist. For example, data interoperability can only be achieved when the used ontologies are interlinked. For that reason, ontologies in the Linked Data cloud can be used, because links between semantic entities that represent the same real-world entity are identified. There are services such as sameAs[25] that provide alignments, for example between the music entities of the BBC Music Ontology[26] and MusicBrainz.[27] Several other alignments have been created by the NoTube project to enable the use of ontologies in the TV domain, such as TV-Anytime.

Another issue is the availability of URIs. An example of unavailability of URIs is the genres of the dutch TV EPG guide: the genres do not have official. Without a URI, one cannot reference it in semantic data, thus URIs have to be created in a standarized fashion, so that when the EPG data is harvested on different computers, the same URI is produced. If not, the data is not interoperable Other entities have many URIs, such as movies. IMDB is a renown movie data repository and IMDB URLs are often used for movie identifiers, such as by Boxee. However, there are alternatives such as Freebase, which also has a URI for virtually every movie. Fortunately, this specific example is solved by including a link to the IMDB counter part of every movie on Freebase.

Besides using semantics, the NoTube recommender system

[21] http://xmlns.notu.be/wi/
[22] http://www.w3.org/TR/skos-reference
[23] http://www.bbc.co.uk/blogs/radiolabs/2009/07/the_music_bore.shtml
[24] http://boxee.tv/
[25] http://sameas.org
[26] http://www.bbc.co.uk/music/
[27] http://musicbrainz.org/

will take contextual parameters into account, such as location and time. For example, somebody who has the habit of watching the news at work does not want to see a soap opera recommended when he/she is at work, even though he/she loves soap operas. The location can be deduced by means of a GPS system that is available in most smartphones or the IP address of the device that is used to access the service.

6. PROGRESS REPORT AND FUTURE WORK

For the NoTube project we have developed several alignments between genre vocabularies such as the TV-Anytime, BBC and YouTube genres. We have developed a fully functioning Beancounter service, which is able to extract data from several social Web services, including LastFM.[28] Other fully functioning services we developed include a named entity recognition service, which identifies DBpedia concepts, and a data warehouse service that provides enriched EPG data from over a hundred TV channels.

Future work includes significantly improving the current simple implementations of the recommender service, set top box software (based on MythTV) and the NoTube Web Service.

7. ACKNOWLEDGMENTS

This work is supported by the EU IP Project NoTube.

8. ADDITIONAL AUTHORS

Lora Aroyo (Vrije Universiteit Amsterdam, email: l.m.aroyo@cs.vu.nl),

9. REFERENCES

[1] R. Baeza-yates and B. Ribeiro-Neto. Modern information retrieval, 1999.

[2] C. Biancalana, A. Micarelli, and C. Squarcella. Nereau: a social approach to query expansion. In *WIDM '08: Proceeding of the 10th ACM workshop on Web information and data management*, pages 95–102, New York, NY, USA, 2008. ACM.

[3] P. Brusilovsky and M. T. Maybury. From adaptive hypermedia to the adaptive web. *Commun. ACM*, 45(5):30–33, 2002.

[4] F. Carmagnola, F. Cena, C. Gena, and I. Torre. A semantic framework for adaptive web-based systems. In P. Bouquet and G. Tummarello, editors, *SWAP*, volume 166 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2005.

[5] T. Finin and D. Drager. GUMS: A General User Modeling System. In *Proceedings of the workshop on Strategic computing natural language of the Human Language Technology Conference*, pages 224–230. Association for Computational Linguistics, May 1986.

[6] M. Minio and C. Tasso. User modeling for information filtering on internet services: exploiting an extended version of the um shell. 1996.

[7] D. Palmisano, M. Minno, D. Brickley, and M. Mostarda. A semantic web based solution for an autocompletion system. January 2009.

[8] P. Pu and L. Chen. Trust building with explanation interfaces. In *IUI '06: Proceedings of the 11th international conference on Intelligent user interfaces*, pages 93–100, New York, NY, USA, 2006. ACM.

[9] G. Salton and M. J. McGill. *Introduction to Modern Information Retrieval*. McGraw-Hill, Inc., New York, NY, USA, 1986.

[10] Y. Wang, N. Stash, L. Aroyo, P. Gorgels, L. Rutledge, and G. Schreiber. Recommendations based on semantically enriched museum collections. *Web Semant.*, 6(4):283–290, 2008.

[28] http://last.fm/

Semantic TV-resources brokering towards future television

Hong Qing Yu
Knowledge Media Institute
The Open University
Milton Keynes, UK, MK7 6AA
Tel: +441908653800
h.q.yu@open.ac.uk

Stefan Dietze
Knowledge Media Institute
The Open University
Milton Keynes, UK, MK7 6AA
Tel: +441908653800
s.dietze@open.ac.uk

Neil Benn
Knowledge Media Institute
The Open University
Milton Keynes, UK, MK7 6AA
Tel: +441908653800
n.j.l.benn@open.ac.uk

ABSTRACT
Nowadays, more and more distributed digital TV and TV-related resources, such as Electronic Personal TV Guide (EPG) data, are published on the Web. To enable applications to access these resources easily, the TV resource data are commonly provided by Web service technologies. The huge variety of data related to the TV domain and the wide range of services that provide it, raise the need to have a broker to discover, select and orchestrate services to satisfy the runtime requirements of applications that invoke these services. The variety of data and the heterogeneous nature of service capabilities make it a challenging domain for automated web-service discovery and composition. To overcome these issues, we propose a two-staged service annotation approach, which is resolved by integrating Linked Services and the IRS-III semantic web services framework, to complete the lifecycle of service annotation, publishing, deploying, discovery, orchestration and dynamic invocation. This approach satisfies both the developer and application's requirements to use Semantic Web Services (SWS) technologies manually and automatically.

Categories and Subject Descriptors
D.2.11 [**Software Architectures**]: Service-oriented architecture (SOA)

General Terms
Design, Experimentation, Human Factors.

Keywords
Semantic Web Services, Semantic Web, Linked Data, Digital TV.

1. INTRODUCTION
With the high demands of Digital TV and IP TV, more and more TV broadcast organizations provide TV-related content as multimedia resources to be accessed through Web technologies [3]. This allows audiences to better interact with the broadcast content and allows them to watch programs via various kinds of media and devices. For example, a user can view Electronic Program Guide (EPG) data via an iPhone, or get personal recommendations via a TV Box. Also, a user can search for a TV program that she/he is interested in and retrieve the actors' detailed information along with the broadcast information of the TV program. To enable end-user applications to access these TV resources easily, the TV resource data are commonly provided by Web service technologies. There are many different service message formats and many different service functionalities available such as video processing services, video transcoding services, EPG related services, etc. Dealing with this level of heterogeneity is a major challenge that raises the need to have a broker to discover, select, mediate, and orchestrate services to satisfy the runtime requirements of the end-user applications. Furthermore, with large-scale availability of TV online resources, such a middleware component is essential for not only sharing the resources but also enabling interoperability between the functionalities that use and process the resources.

In this paper, we propose and implement a Semantic TV Resource Broker (STRB) based on a Service Oriented Architecture (SOA) approach in combination with Semantic Web Services standards. In other words, the functionalities are deployed as Web services with semantic metadata. Here we refer to a Web service as any kind of software functionality that is accessible through HTTP, varying from REST-based APIs to SOAP interfaces.

To achieve this goal, we need to address some important challenges:

- *Developing an easy approach that allows service developers to annotate semantic metadata for their TV-related services.* Currently, semantic annotations are mainly based on WSMO [17] and OWL-S [8]. However, there is a gap between developer's knowledge about these ontologies and tools for supporting developers. In addition, the complexity of the WSMO and OWL-S standards impedes the adoption by the developers' community to use them in large-scale applications.
- *Enabling both developers and applications to seamlessly interact and align the available semantic metadata provided by developers.* This challenge is to bridge the somewhat conflicting requirements. A developer needs an easy-to-understand and human-readable description of the functionalities of the services. In contrast, in order for applications to interact and integrate services, they require a machine-oriented environment to dynamically work with services and their metadata. Therefore, the service metadata annotations should have two different levels of interoperability.

The contributions of our approach in meeting the above challenges are:

- We implement a two-staged approach for semantically annotating the services in the STRB. The two stages are (1) allowing developers to annotate and publish the services by using the *Linked Services* [2] approach based on lightweight RDF annotations through a Web form; (2) the Linked Services

RDF annotations feed into the IRS-III [4] semantic execution environment to semantically deploy the service. The term Linked Services is used to describe the fact that the semantic service annotations using this approach are much easier to produce (than say those based on WSMO or OWL-S) and can be populated with references to widely established Linked Data vocabularies. Furthermore, they address a much wider audience and allow even non-SWS experts and lay people to describe and annotate services.

- With these two levels of annotations, the developer can manually discover and select services by simply using SPARQL [16] queries to develop applications on top of the STRB. Meanwhile, applications can invoke atomic services or orchestrated services through the IRS-III semantic execution environment that can automatically discover, orchestrate and invoke the available Web services.

This paper is organized as follows: Section 2 introduces the background information about the NoTube project and related terminologies and technologies. Section 3 shows two realistic use cases that present the main requirements of the STRB. Section 4 illustrates the architecture of STRB and the detailed implementation of the prototype. Finally, the conclusion and future work are discussed in Section 5.

2. BACKGROUND

The STRB is defined as an important middleware component for the NoTube project with the purpose of automatically finding, combining and invoking relevant Web Services based on goals specified by the NoTube application developers. The ultimate goal of this project is to develop a flexible/adaptive end-to-end architecture, based on semantic technologies, for personalized creation, distribution and consumption of TV content. The project takes a user-centric approach by investigating the fundamental aspects of consumers' content-customization needs, interaction requirements and entertainment wishes, which will shape the future of the television experience. Figure 1 shows the overall NoTube environment, containing four conceptual layers of service, broker (or called control), application (or called view) and screen. The broker is centrally located in the architecture because it is responsible for the communication between applications and services.

At this moment, we have collected more than 40 existing services relevant to the TV domain. The functionality of the services contains EPG services, context logging services (e.g. user profiling service), enrichment services (getting richer information about a certain TV resource from multiple resources), social network services (like interfaces to the functionality of. Twitter or Facebook), and recommendation services.

The technology used to develop the broker is based on SWS. The broker uses a repository of SWS in order to perform its functionalities. SWS are Web Services enriched with ontological descriptions of Web services in terms of their capabilities, interfaces and non-functional properties. SWS technologies aim at automatic discovery, selection and orchestration of distributed services for a particularly expressed user's request/goal. The SWS approach utilizes both standard Web service technology such as SOAP [13], UDDI [14] and WSDL [15] and more lightweight approaches such as REST or XML-RPC.

Figure 1--NoTube overall architecture.

The current efforts of the SWS research community have resulted in reference ontologies, such as OWL-S, WSMO and SAWSDL[1] as well as comprehensive frameworks to demonstrate the SWS approach (i.e. DIP[2] project). Whereas WSMO is intended to enable fully automated service matchmaking based on comprehensive semantic specifications of service capabilities, recent derivations of WSMO, like WSMO-Lite [11], and MicroWSMO [5] lightweight service descriptions based on RDF.

Most recently, the Linked Services concept has been proposed based on Linked Data principles. Linked Data is a way to publish data on the Web in order for machines to automatically derive the meaning of the data. The Linked Data cloud contains a rich variety of alignments between external data, which makes it possible to create services that make use of relevant combinations [1]. We implement the STRB by integrating the available Linked data with the IRS-III semantic web services framework.

3. USE CASES AND CHALLENGES

In order to explain the role and functionalities of the STRB, we select and illustrate two use cases that are driven by the TV broadcast industry partners within the EU NoTube project.

3.1 Personalized Semantic News

[1] http://www.w3.org/2002/ws/sawsdl

[2] http://dip.semanticweb.org/

The Personalized Semantic News use case describes how a user acquires news items from generic broadcast streams and obtains additional enriched news information by using a set of personalised news related services through the NoTube platform. The NoTube platform understands the meaning of video news items and the physical context in which news items are going to be shown. Based on this, the platform will apply criteria for matching and filtering the user profile and preferences to match the available news items. Figure 2 shows a possible scenario where a user asks his/her context-aware news-agent to search interesting news when he/she is using an iPhone and travelling by bus. He/she registered his/her profile to the agent and he/she prefers to use English and is generally interested in sports. The agent will invoke the STRB to get the interesting news data by discovering, selecting and invoking the suitable news services that match the user's context.

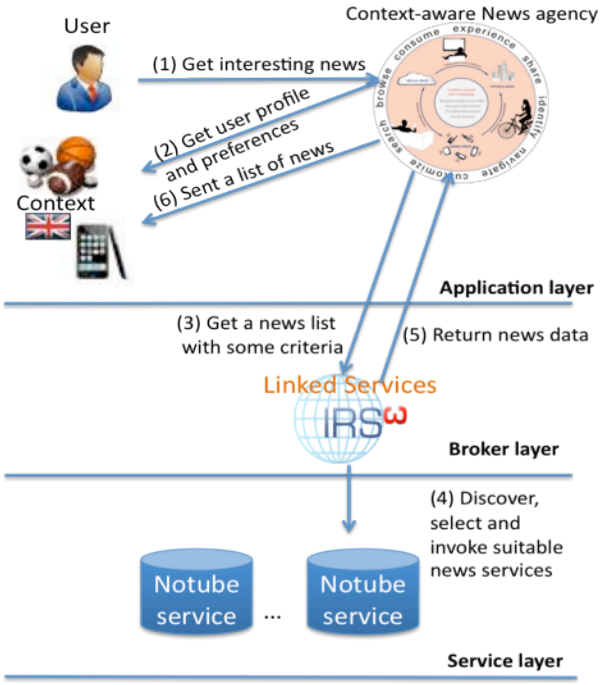


Figure 2. Personalized semantic news use case.

There are some challenges that need to be addressed for achieving this use case. The first is how services should publish their metadata about functional and non-functional properties in order to be discovered, selected, orchestrated and invoked. The second challenge is what ontology should be used to describe the TV resources services. The third challenge is what mechanism should be applied to automatically discover, select, orchestrate and invoke the suitable services for the application at runtime.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
EuroITV NoTube Workshop, June 9, 2010, Tampere, Finland.
Copyright 2010 ACM 1-58113-000-0/00/0010…$10.00.

3.2 Personalized TV Guide

This use case allows a user to send a request for getting EPG data with program recommendations and additional related information gained from Internet resources that are provided by a set of TV program enrichment services. The recommendations should be based on the context of the user, such as user activities, languages and personal interests. This information is stored by a User Profiling service. Thus, the core services that underpin this scenario are User Profiling services, EPG services, TV program recommendation services, and content enrichment services.

Unlike the first use case, the Personalized TV Guide scenario requires the broker to orchestrate a group of services according to an orchestration process shown in Figure 3. The orchestration process does not combine any concrete services at the beginning and only assigns services at runtime, which is the main difference contracting to current WSBPEL[7] technology (e.g. activeBPEL and ODE). Thus, the additional challenge raised by this use case is how to dynamically orchestrate different services throughout discovery, selecting and invoking steps.

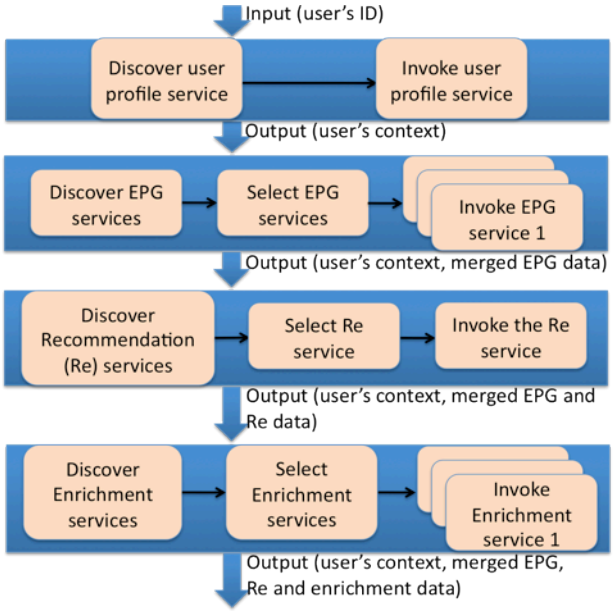


Figure 3. Personalized TV Guide use case orchestration.

4. STRB ARCHITECTURE

Our STRB solution integrates two semantic web technologies although they are currently still under development for improvement, namely, Linked Services and the IRS-III framework. Figure 4 shows the core development workflow in which the STRB operates. That is:

- Service providers annotate, register and publish their services into the Linked Services RDF repository that adapts Linked Data principles for linking services to their functional and non-functional annotations and other services in order to dynamically discover services. It is as simple as searching a web page and automatically selecting services based on required properties described as SPARQL query.
- After service registration, the services repository will feed the new functional annotations into the IRS-III framework. The

registered service can be discovered, selected, orchestrated by the IRS-III semantic execution environment.

- The Application Developer can request invocation URIs from the STRB for the required services to develop NoTube applications.

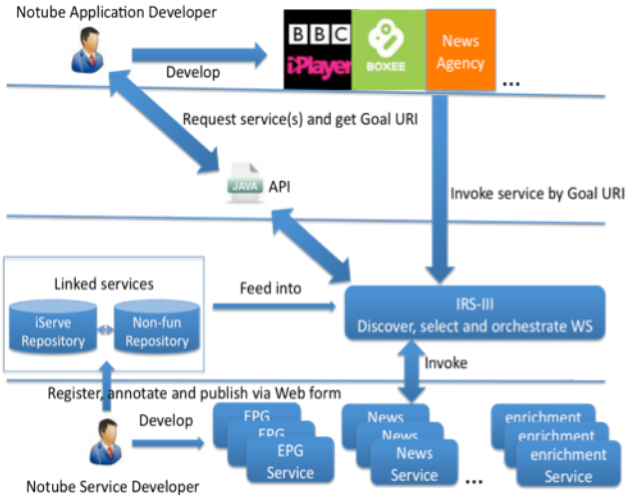


Figure 4. NoTube services development workflow.

4.1 Linked Services

The idea of Linked Services is inspired by the Linked Data movement. Linked Data is a way to publish data on the Web in order for machines to understand the explicit meaning of the data. The data is linked to other external data sets, and can in turn be linked from external data sets [9]. In this way, the data can be found and operated directly by machines.

Based on a similar idea, we publish "linked services" with their semantic descriptions on the Web. The basic principles are:

- Using WSMO-Lite and MicroWSMO as functional semantic description schema and using a number of domain ontologies as non-functional semantic description schema.
- Representing and persisting the semantic descriptions as RDF data stored in a Sesame RDF database. The database includes two divided RDF repositories: (1) iServe [10] for storing functional descriptions, such as invocation endpoints, input messages and output messages and so on; (2) non-funServe for storing non-functional descriptions, such as IRS-III goal URIs, QoS and keywords and so on.
- Allowing service providers to annotate publish their services using a web-based form UI (see Figure 5). Figure 6 shows an example of linked data for semantically describing a service after annotated and published by provider via the web-based form.
- Using a Similarity-based Conceptual Space approach proposed in [2] for semantic web service discovery and selection.

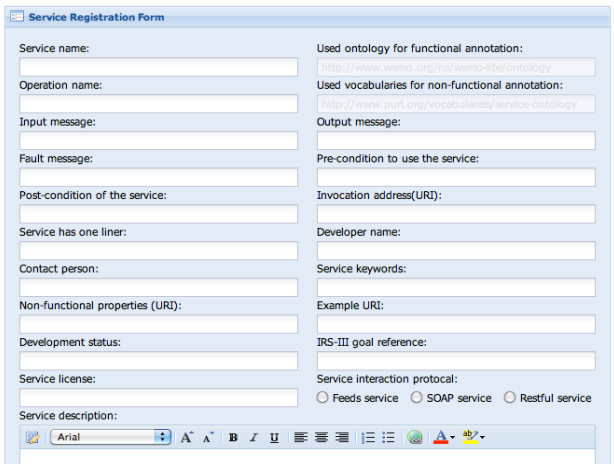


Figure 5. Linked services publishing user interface.

Subject	Predicate	Object
<http://www.notube.tv /services#ExtractGenreData>	rdf:type	<http:/
<http://www.notube.tv /services#ExtractGenreData>	<http://www.w3.org/ns/sawsdl#modelReference>	a:Video
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasliner>	"Get Ge
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasDescription>	""
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasDeveloper>	"IRT"
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasContactPerson>	"R. Zim
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasKeywords>	<http:/
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasCreationDate>	<http:/
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasQoS>	<http:/
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasProtocal>	<http:/
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasStatus>	<http:/
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasExample>	""
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasLicense>	""
<http://www.notube.tv /services#ExtractGenreData>	<http://www.purl.org/vocabularies/service-ontology#hasGoal>	<http:/
<http://www.notube.tv /services#ExtractGenreData>	owl:sameAs	<http:/ eadd-4
<http://www.notube.tv /services#ExtractGenreData>	<http://www.wsmo.org/ns/wsmo-lite#hasOperation>	<http:/

Figure 6. A RDF data example of a Linked service annotation.

4.2 The IRS-III Framework

IRS-III is a Semantic Web Service execution environment. It acts as "broker" – mediating between the goals of a client and relevant services that are deployed on the Web. IRS-III adopts the WSMO conceptual model of services, the ultimate aim of which is to be able to provide unambiguous models of services with a well-defined semantics, which can then be interpreted by a reasoner to enable automatic discovery, selection, composition, mediation, execution, and monitoring of services.

At runtime, IRS-III automatically discovers and invokes Web services suitable for a given client request, formulated as a goal instance.

The Broker's functionalities are exposed through an HTTP-based REST API, which applications use to interact with the Broker. Primary among the functionalities of the Broker is the 'achieve-goal' method. In the IRS approach, Web services are semantically described and associated to "goal" representations. Goals are then exposed through the Broker's REST interface to allow them to be "achieved" by the client application. We have also implemented a Java API to the IRS-III Broker, details of which can found in **Error! Reference source not found.**.

5. CONCLUSION AND FUTURE WORK

In this paper, we described the implementation of a Semantic TV Resources Broker. The broker allows application developers within the TV domain to express application goals in human-readable format, and discovers, integrates and invokes the relevant services in order to fulfill these goals.

The technologies that we used are: (1) Linked service repositories for publishing services and annotating functional and non-functional properties; (2) the IRS-III framework including the Java API for service logical annotation based on WSMO, service invocation and orchestration through achieving-goal interface.

At present our work is at an early stage and we still need to evaluate the entire two-stage approach. In the future, we will conduct this evaluation by comparing our work to other related frameworks, for example, the Dino framework of Dynamic and Adaptive Composition of Autonomous Services [6], to analyze the efficiency, adaptability, and scalability of the different approaches.

Other future work includes elaboration on more advanced selection mechanisms, e.g. the LSP-based selection method [18], into the implementation to test and compare the selection and composition performance.

6. REFERENCES

[1] Bizer, C., T. Heath, et al. (2009). "Linked data - The Story So Far." Special Issue on Linked data, International Journal on Semantic Web and Information Systems (IJSWIS).

[2] Dietze, S., Benn, N., Domingue, J., Conconi, A., and Cattaneo, F.: "Interoperable Multimedia Metadata through Similarity-based Semantic Web Service Discovery". In Proceedings of 4th International Conference on Semantic and Digital Media Technologies (SAMT '09), 2--4 December 2009, Graz, Austria.

[3] Dietze, S., and Domingue, J. (2009) Towards Context-aware Multimedia Processing through Semantic Web Services, in Proceedings of the seventh european conference on European interactive television conference (EuroITV 2009), ACM, pp. 129-132, Leuven, Belgium.

[4] Domingue, J., Cabral, L., Galizia, S., Tanasescu, V., Gugliotta, A., Norton, B., Tanasescu, V., Pedrinaci, C. (2008): IRS-III: A Broker-based approach for Semantic Web Services. Journal of Web Semantics pp. 109—132. Elsevier Science Publishers.

[5] Maleshkova, M., Pedrinaci, C., and Domingue, J. (2009). Supporting the creation of semantic RESTful service descriptions. In workshop: Service Matchmaking and Resource Retrieval in the Semantic Web (SMR2) at 8th International Semantic Web Conference.

[6] A. Mukhija, A. Dingwall-Smith, and D. S. Rosenblum, 2007. QoS-Aware Service Composition in Dino. In Proceedings of the Fifth European Conference on Web Services (November 26 - 28, 2007). ECOWS. IEEE Computer Society, pp. 3-12.

[7] Organization for the Advancement of Structured Information Standards (OASIS), WSBPEL: Web Service Business Process Execution Language 2.0, (2007). (http://docs.oasis-open.org/wsbpel/2.0/OS/wsbpel-v2.0-OS.html)

[8] OWL-S Coalition: OWL-S 1.1 release. (2004). http://www.daml.org/services/owl-s/1.1/

[9] Pedrinaci, C., Domingue, J., and Reto Krummenacher (2010) Services and the Web of Data: An Unexploited Symbiosis, Workshop: Linked AI: AAAI Spring Symposium "Linked data Meets Artificial Intelligence".

[10] Carlos Pedrinaci, Dave Lambert, Maria Maleshkova, Dong Liu, John Domingue, and Reto Krummenacher. Adaptive Service Binding with Lightweight Semantic Web Services. In Service Engineering: European Research Results (S. Dustdar and F. Li eds.). Springer (2010).

[11] Vitvar, T., Kopecky, J., Viskova, J., and Fensel, D. (2008). WSMO-Lite annotations for Web services. In Hauswirth, M., Koubarakis, M., and Bechhofer, S. eds., Proceedings of the 5th European Semantic Web Conference, LNCS Berlin, Heidelberg: Springer-Verlag.

[12] Wagner, M., Kellerer, W. 2004. Web services selection for distributed composition of multimedia content, Proceedings of the 12th annual ACM international conference on Multimedia, October 10-16, 2004, New York, NY, USA.

[13] World Wide Web Consortium, W3C: Simple Object Access Protocol, SOAP, Version 1.2 Part 0: Primer, (2003). (http://www.w3.org/TR/soap12-part0/).

[14] World Wide Web Consortium, W3C: Universal Description, Discovery and Integration: UDDI Spec Technical Committee Specification v. 3.0, (2003). http://uddi.org/pubs/uddi-v3.0.1-20031014.htm).

[15] World Wide Web Consortium, W3C: WSDL: Web services Description Language (WSDL) 1.1, (2001). (http://www.w3.org/TR/2001/NOTE-wsdl20010315)

[16] World Wide Web Consortium,W3C Recommendation, SPARQL query language for RDF, 2008, (http://www.w3.org/TR/rdf-sparql-query/).

[17] WSMO Working Group (2004), D2v1.0: Web service Modeling Ontology (WSMO). WSMO Working Draft, (2004). (http://www.wsmo.org/2004/d2/v1.0/).

[18] Yu, H.Q., Reiff-Marganiec, S., "A Method for Automated Web Service Selection", services, pp. 513-520, 2008 IEEE Congress on Services - Part I, 2008.

Integrating social networks in an IPTV recommender system

Cerdeira, Frederico	Almeida, Pedro	Abreu, Jorge	Almeida, Pedro L.	Cardoso, Bernardo
University of Aveiro	University of Aveiro	University of Aveiro	University of Aveiro	PT Inovação
3810-193 Aveiro - Portugal	3810-193 Aveiro - Portugal	3810-193 Aveiro - Portugal	3810-193 Aveiro - Portugal	3810-106 Aveiro - Portugal
+351 234 370389	+351 234 370389	+351 234 370389	+351 234 370389	+351 234 403200
frederico.cerdeira@ua.pt	almeida@ua.pt	jfa@ua.pt	pedroalmeida@ua.pt	bernardo@ptinovacao.pt

ABSTRACT

The proliferation of multiple channels in the actual offer of cable and IPTV services is bringing challenges for users when selecting which program to see. Recommender systems are being introduced with the goal of facilitating the viewers' choices and promote TV content consumption. Different technical approaches to these systems are supported either on automatic algorithms or social data and activities analysis.

This paper reports on the first development phase of the Pétala system - a recommender system for an IPTV Telco - that aims to improve content recommendation by integrating social networks. Pétala provides the ability to use the system both in the Set-top box or as a web widget integrated in a social network like Facebook.

Categories and Subject Descriptors

H5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Prototyping, Screen design, interaction styles.*

General Terms

Algorithms, Design, Experimentation.

Keywords

Social networks, IPTV, recommendation, EPG, social TV.

1. INTRODUCTION

Social TV as a part of an interactive television distribution service may refer to the ability to communicate between two or more users at distance. For that different communication services may be used, like open audio channels, IM services (*instant messaging*) or other text based services complemented with information on channels/programs being viewed by each user.

However, even with this type of awareness, the wide variety of TV channels offered by IPTV or cable TV infrastructures opens a scenario where content recommendation based on the users' preferences is increasingly justified. Actually, this scenario contributes not only to the fragmentation of TV's audiences but also to a possible feeling of disorder since the user may get lost when he wants to choose a TV channel or program.

Content recommender systems may appear as a way to filter this wide TV offer allowing viewers to get recommendations on the programs related with their tastes and interests. Recommender Systems (RS) may base their information in the profile of the user along with the information retrieved by the activities in Web 2.0 social networks. These may act as triggers to perform content recommendation in a TV environment.

The Pétala system appears as a research prototype that aims to explore ways to integrate social networks (their features and information) in an IPTV environment.

2. RELATED WORK

A typical RS can provide recommendations to the user according to his profile, predicting his personal likes, typically according to his pattern of TV consumption or the ratings that he has assigned to similar content. In this way a RS will typically present items that have high rating levels [6].

RS are usually based in *Collaborative Filtering algorithms*, allowing, for example, that the system suggests an item through the matching of the user profile with the profiles of other users. In this kind of systems, data from a multiple-value rating (for example, in a 1 to 5 scale) is considered to evaluate items/contents [8]. Consequently, the system will be able to recommend similar items to the ones previously approved by the user or other users. Another strategy relies in *Content-Based* methods. These systems are based on semantic analyses, comparing the user items with a list of possible related ones. These strategies are typically applied in RS that are based in a complex method using algorithmic techniques that rely in several sources of data. Yet, content recommendation can be based in community (data) source [3], where socially networked users play an active role as content recommendation agents.

Content recommendation should be understood as a way to suggest content contributing for the minimization of the information overload typically associated with the EPG [7] [9]. Brokens [5] also argues that a recommendation should be trusted by the user along with the control over what is recommended. Several researchers and players, in the last years, have developed solutions of content recommendation for interactive TV, IPTV and CATV applications. Some relevant examples are the case of TiVO[1] or the ConnecTV project [4].

Recently, in the latest Microsoft's MediaRoom Application Contest[2], the Accedo BroadBand[3] introduced an application based in a community where users can see *tweets* and Wikipedia

[1] http://www.tivo.com/ (December 2009)

[2] http://www.microsoft.com/Mediaroom/ContestWinners.aspx (December 2009)

[3] http://accedobroadband.com (December 2009)

information related with TV programs. Along with this enhanced information, users can also vote on content/programs.

Most of the research projects focus on user's preferences and their ranking of movies and TV programs as the main source of information. This method can be in some away disadvantageous because it requires user time, will and effort [10]. Most of the research projects are applied to a one-to-one scenario in an individual use of the RS. Providing content recommendations for multi-user environments is also a concern. The work developed in the project "unobtrusive context-aware recommender system" [10] presents interesting results because it relies in a multi-user environment and the used methods do not force users to provide pre-information such as preferences or interests. In the *WeOnTV* project [1], users can know what others are viewing and they can offer direct recommendations to other in a buddy list. However, this is a social approach to content recommendation rather than an automatic RS.

Collabora TV [9] a research project that explores social activities supported by asynchronous communication, allows viewers to communicate through text annotations superimposed on programs. The viewers can access a list of shows that their buddies have been watching or track a list of *Popular Shows*. This is a different method of recommendation.

Another important related project is the *Television Meets Facebook* [2]. This project aims to improve the experience of TV consumption through the convergence between TV and an interactive social network. This application relies in connecting the STBs to the *Facebook* API to suggest TV content that friends like (based on their ratings). It also allows automatic content recording.

In the Portuguese commercial scenario, Telcos are offering different IPTV products and solutions. The most popular is MEO[4] from Portugal Telecom. *Sonaecom*, a player in the telecommunication market, offers the *Clix* TV solution. Besides VOD, DVR and Restart TV[5] services, *Clix* also offers Digital TV with a content recommendation service. The application suggests movies and TV shows based on the user preferences (however these need to be manually introduced).

Like the Collabora TV and WeOnTV projects, the **Pétala System** uses social methods to provide recommendations, but instead of having a specific buddy list, the **Pétala System** uses a social network like Facebook, to retrieve recommendations from the user's friends. In the **Pétala System** the recommendations can be retrieved directly from the friends (like a direct message) or, in a social approach, can be displayed as an average rating in the EPG. Since the RS services offer in the Portuguese Telcos is almost insignificant, it would be interesting to explore content recommendation based on social networks and realize how to implement a project like the **Pétala System** in a commercial IPTV solution.

[4] http://meo.pt (April 2010)

[5] Feature that lets you go back to the beginning of a program that has already started

3. THE PÉTALA SYSTEM

3.1 Goals

The main goal of the project is the research and development of strategies for the integration of Web 2.0 social paradigms and networks to achieve a content recommendation system over an IPTV environment.

It was defined as a primary goal the development of a RS based not only on inference algorithms but mainly in information from social networks, namely the users' friends activity (votes and recommendations) in these social networks.

This **Pétala System** includes two main client applications: i) an application in the STB; ii) a widget type application integrated in social networks like Facebook or Hi-5. For both modules a special EPG was prepared enriched with the ability to recommend/vote a program while watching it (if in TV) or while using Facebook or Hi-5.

3.2 Application Modules

The system is organized in three interconnected modules.

The Web Widget module consists of an application integrated into Facebook and Hi-5. It includes an enhanced EPG containing the programming grid from MEO[4] including a representation of reckoning rates applied to each program. This module includes a rating functionality where users can classify TV programs. It also allows users to send direct recommendations to other users (in a format of a pre-set text message). On the other hand it allows reading direct recommendations from others, to see how a program is rated or to view tops of the most rated programs.

The TV-STB module is a prototype developed in the MediaRoom – Presentation Framework Middleware. Like the Web Widget module, it also relies in an EPG with information of the reckoning rates associated to each program, the possibility to rate a program and to send direct recommendation to friends. The top of most rated programs is also available.

Finally, a Server-Side module supports both client applications. It consists of a layer of web-services informed by the system's database. These web-services act as a (normalized) bridge between the Web Widget and the TV-STB application.

4. THE DEVELOPMENT PROCESS

4.1 System Architecture

For supporting both client applications, Web and TV, a set of web services implemented in a cloud computing paradigm is used. Figure 1 represents all the components involved: a *Database Server* hosts the main SQL database that stores all the users' interaction activity (ratings and direct recommendations); the *Web Server* is responsible for all ASP.NET modules (TV-STB APP, Web Widget and Pétala Web Service) relying in Microsoft Internet Information Services 6.0; the *Social Data* server, hosting the *Facebook API Platform*[6], is accessed to retrieve information from that social network; and the *SAPO EPG API*[7] is the source of information for the (MEO) EPG. This API is used to populate the enhanced EPG in the TV-STB and Web Widget applications.

[6] http://developers.facebook.com (April 2010)

[7] http://services.telecom.pt/ (April 2010)

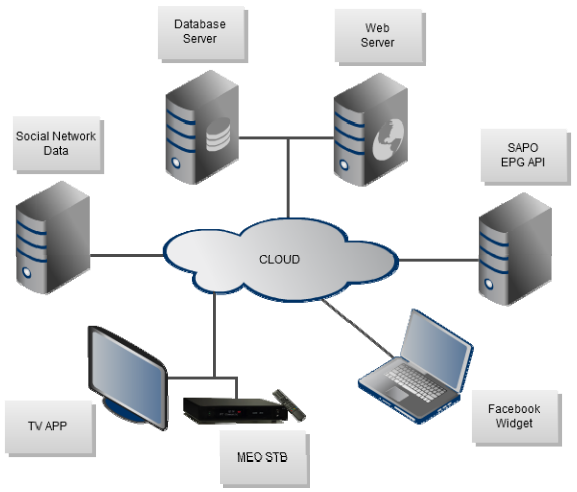


Figure 1. System Architecture

4.2 Server-Side Module

The Web Service Pétala, relying on a XML model developed in ASP.NET – C#, assures the communication with the server and its database. Using this module the client application does not directly operates with the server or the database, minimizing data inconsistency and bugs. It also adds an extra level of safety in the access to the system.

It offers several reading and writing methods like returning buddy lists, reckoning rates or direct messages and input operations like rating and sending direct messages.

One of the methods returns a list of the "TOP" ranked programs. In this list every program has a representation of its arithmetic average, based on all rates. It is also possible to set a social filter on this ranking, returning the most rated programs (and its average) from one specific social network. The user may also decide to retrieve the ranking based on all users' data or only on his friends' data. Despite some limitations, these filters introduce some level of personalization to Pétala.

4.3 Web Widget

This module includes three informative panels: EPG, TOP and INBOX (see Figure 2).

Figure 2. EPG Panel

The EPG Panel shows the program guide for the next two hours but can retrieve information for a week period. The TV programs are identified by its name and the rating information is overlaid in the same area.

By clicking in a program a floating panel presents two options: rate the program or recommend it to a buddy. When rating the user can choose between a 1 to 5 scale (which will then contribute to the reckoning of that program). If he chooses to recommend he can select a buddy from his contact list and the recommendation will go as a direct message to the buddy's INBOX.

Figure 3. TOP Panel

The Top Panel shows the scores, represented by stars, of all rated programs in the EPG, being possible to the user access its schedule information. The user can also filter the results according to his preferences: the first option toggles the reckoning of the TOP values taking in consideration inputs from the entire set of users of the recommendation system or only from his buddies; it is also possible to split the ratings from different social networks; finally the user can also choose the TV programs of one TV channel by selecting his favorite channel in the filter. These filters are cumulative and produce changes in the final calculation of the "rating" of each program since the sample collection will change every time the user changes the filtering options.

The INBOX panel is a small mailbox of direct recommendations from other users, displaying the sender's name, the recommended program and the sending date. For further information about the transmission of the TV program there is a button that triggers a floating panel with the schedule information.

4.4 TV-STB APP

This application was originally developed in a simulator environment - ADK (Application Development Kit) - Version 2.1 (June 2009), followed with the deployment in a STB equipped with a MPF build: Model Cisco KMM3010.

The TV-STB prototype comprises tree modules:

Pétala EPG – Enhanced EPG where it is possible to rate and recommend content;

Received Recommendations – an area where the user can access recommendations sent from other users.

Rating Tops – an area displaying the most high rated programs.

Figure 4. Rate or Recommend options

The Pétala EPG (Figure 4) consists of a program guide in its minimal style, identical to the existing commercial MEO solution. By consulting a program's information the viewer can: i) see information about the rating average retrieved from social networks (the users' friends' votes) and other Pétala users; ii) rate the selected program (figure 5). The viewer can assign a value (from 1 to 5 stars), see how many people rated that program and its reckoning;

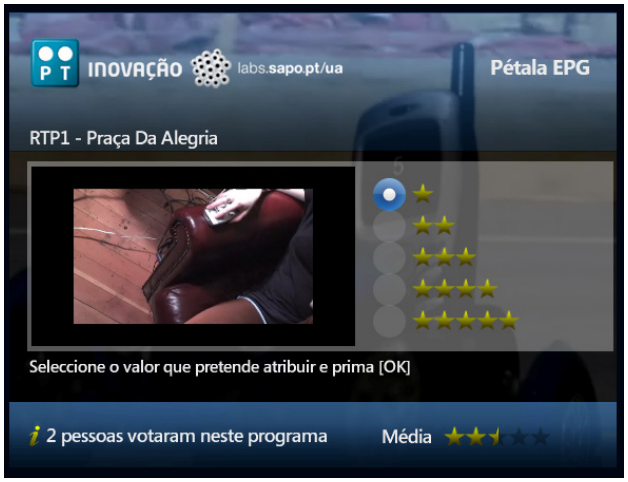


Figure 5. Rating area

iii) recommend it to a friend through a direct message – the viewer just needs to select the friend from his contact list.

In the STB the user can also read program recommendations sent by his friends (either from the Web Widget or from TV-STB APP). These recommendations are complemented with information on when the program will be broadcasted again.

5. REMARKS AND FUTURE WORK

Although in an early development phase the work done so far has allowed the team to start exploring means of integrating social networks data in a recommender system. Despite some positive feedback from users in informal public demonstrations, further and more structured evaluation is needed and will be done in near future.

As a prototype, the Pétala system only supports a part of the features that the research team expects to deliver in a future version of the RS. Considering future developments an improved integration in the middleware is aimed. One way to achieve it could be by means of integrating these features in a more complete social TV application like WeOnTV.

Another planned feature includes the ability to provide emotional mood as a parameter for the recommendation.

Another possible feature relates with the ability to have a visual representation of which content is being/was consumed based on different parameters like: geographical area, age or genre.

6. ACKNOWLEDGMENTS

Our thanks to the research fellow P. Ricardo. Thanks also to the PTIn and labs.sapo/UA teams.

7. REFERENCES

[1] Abreu, J.F., Almeida, P., Pinto, R. and Nobre, V. Implementation of social features over regular IPTV stb *Proceedings of the seventh european conference on European interactive television conference*, ACM, Leuven, Belgium, 2009, 29-32.

[2] Baca, M. and Holtzman, H., Television meets Facebook: Social Networks through Consumer Electronics. in, (2008).

[3] Bernhaupt, R., Wilfinger, D., Weiss, A. and Tscheligi, M. An Ethnographic Study on Recommendations in the Living Room: Implications for the Design of iTV Recommender Systems. *Changing Television Environments*. 92-101.

[4] Boertjes, E., Klok, J. and Schultz, S. ConnecTV: Results of the Field Trial. *Adjunct Proc. EuroITV*.

[5] Broekens, J. Want a Recommendation? Ask my Digital Personality! *6th European Conference, EuroITV 2008*, Salzburg, Austria, 2008.

[6] Cosley, D., Lam, S., Albert, I., Konstan, J. and Riedl, J., Is seeing believing?: how recommender system interfaces affect users' opinions. in, (2003), ACM New York, NY, USA, 585-592.

[7] Lee, H., Ferguson, P., Gurrin, C., Smeaton, A., O'Connor, N. and Park, H., Balancing the power of multimedia information retrieval and usability in designing interactive tv. in, (2008), ACM, 105-114.

[8] McLaughlin, M.R. and Herlocker, J.L. A collaborative filtering algorithm and evaluation metric that accurately model the user experience *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, ACM, Sheffield, United Kingdom, 2004, 329-336.

[9] Nathan, M., Harrison, C., Yarosh, S., Terveen, L., Stead, L. and Amento, B. CollaboraTV: making television viewing social again *Proceeding of the 1st international conference on Designing interactive user experiences for TV and video*, ACM, Silicon Valley, California, USA, 2008, 85-94.

[10] Vildjiounaite, E., Kyllönen, V., Hannula, T. and Alahuhta, P. Unobtrusive dynamic modelling of TV program preferences in a household. *Changing Television Environments*. 82-91.

Digital Video Interaction to support frontal lecturing in multiple platforms of internet based communication

Francesco Epifania
Politecnico di Milano
Dept. of Electronics and Information
Piazza Leonardo da Vinci 32 – 20133
Milano (Italy)
+39 02 2399 9627
Epifania@elet.polimi.it

Giorgio Valle
Università degli Studi di Milano
Dipt. of Computer Science
Via Comelico 39
Milano Italy
02 50316357
Giorgio.valle@unimi.it

ABSTRACT
This paper reports on an active role played by the students of the Digital Communication degree at the School of Sciences of the University of Milan, Italy, to develop e-learning supports both in their university and in the high schools where they studied previously. Creating and maintaining portal websites, discussion forums, learning objects to be exchanged, and more recently, an interactive video-journal that's accessible on mobile phones and WEB-TV, have enhanced the learning process both in the University of Milan and in the aforementioned high schools, through the use of new digital technologies in order to supplement the traditional teaching methods in the classroom, that are bound to the dimensions of space (location) and time. Explanations and feedback regarding the lectures in class are made available through videos uploaded to our WEBCEN server in the LCAD University Research Laboratory, while teaching support is implemented by a community effort which involves teachers as well. Results achieved are measured through the reduction in the amount of insufficient marks among the students at the end of each Academic Year. Interactive Digital Television (IDTV) has opened new frontiers for the diffusion of e-learning support platforms, and when internet connection and internet-based content will also be available on IDTV networks starting from the Spring of 2009, the amount of interaction will reach a new and more significant dimension. Hypertexts will also play a key role in the reachability and fruition of our e-learning support and Interactive VideoJournal applications in multiple platforms with internet connection capability.

General Terms Design, Experimentation

Keywords
Digital, Interaction, video, teaching, Digital Communication, web Tv, social network, Youtube, Facebook, hypermedia, hypertext,

1. INTRODUCTION
An innovative e-learning project was launched a decade ago with the aim to develop a fully functional framework based on the internet to support both the interaction between students and professors, and the distance learning for students who could not attend the lessons based on frontal lecture which requires their presence in the classroom. This framework, known as the Didattica Webcentrica Project (Web-based Learning Aid, www.webcen.unimi.it) still supplements teaching in the Computer Science curricula at the University of Milan[1], as described at the World Conference on E- Learning in Corporate, Government, Healthcare, and Higher Education, 2007 [2].

The Didattica Webcentrica [3] Project allowed us to address specific issues, such as:

- to personalize educational offers, taking into account the students' differences in the background knowledge required by the four of first-level degrees (bachelor's classes 26, 14 and 23/s by Italian law requirement) and the two second-level degrees in the Computer Science curriculum;
- to make educational offerings available to students who can not regularly attend lectures in classrooms, because of part-time jobs or commuting difficulties for far-residing students;
- to enhance higher ICT practice among students, while helping them to obtain better grades.
- This paper investigates video-based interaction as a tool designed to:
- enrich the educational process, by the wide availability of communication equipment such as mobile phones, MP3, and iPod;
- reduce "blackboard" dependency by leveraging on technologies familiar among people, which allow viewing videos on a laptop screen or top-level mobile telephones;
- stress active role plaid by students, through their familiarity with documenting events by digital video-recordings (possibly made public later by uploading to video hosting portals such as YouTube for reaching a world-wide audience) or contributing articles to

magazines/journals promoted by the their high-school or University under a professor's supervision.

The analysis of the new approaches of young people to the communication means (such as new media, mobile access, advanced use of the Internet), first described at an initial stage in our paper at the World Conference on Educational Multimedia, Hypermedia and Telecommunications 2008, Vienna, June 30-July 4 2008 [4], encouraged us to explore other means to improve the efficiency of the learning practices and to enlarge the target audience.

The focus of the Interactive VideoJournal on high school students has two motivations:

- to involve them in an entertaining learning process, with topics that are related to their lifestyle and preferences;
- to get them closer to the university environment, through cooperation with the university students.

It should also be taken into consideration that, based on a research regarding the students' grades in the high schools of the Lombardy Region in Italy, students reached an average of only 60 percent of sufficient marks in the first semester which ended in January 2008, with major difficulties reported particularly in scientific disciplines such as Mathematics and Physics. Even though Italy ranks high in terms of the financial investment in education, a recent EC study comparing 40 nations ranked the performance of the Italian high schools in 37th position. The University of Milan is the third largest university in Italy, with a student population of around 75,000 which grew steadily in the last decade. Our Faculty of Mathematics, Physics, and Natural Sciences is the fourth largest faculty of the university, after the faculties of Medicine, Literature, and Political Sciences. Therefore, finding a way to enhance the quality of incoming new students has become a must, while it is obvious that no university in Italy refuses new students, since the main component of the state's financial support to them is based on the number of enrolled students.

The next paragraphs of this paper will address: the VideoJournal project and VGG (the Interactive VideoGiornaleGiovanile), the experiments with the students of the Law School regarding the VG Servizi Giuridici, the VG Learning environment, content data and metadata management, Pesce d'Aprile Jokes, the VG experimented in the trade fairs LAMIERA and Bi-MU covering the mechanics industry, now to be expended into tourism management, and the promotion of musical & gastronomical events.

2. THE VIDEOJOURNAL PROJECT

The Osservatorio Permanente Giovani Editori, a branch of the FIEG Italian Federation of Newspaper Editors, operates since 1997 a Giornale in Classe Project, www.osservatorionline.it, where a copy of an influential nation-wide newspaper together with a copy of the leading local newspaper are read in the high schools' classrooms for stimulating a debate, lead by a teacher, on specific articles. It is a very successful project, today involving over 1.5 million high school students, expanding in 2007 into 22 Italian universities as well. Financial support comes from bank foundations and major owners of the banks' stocks, while newspapers provide press coverage and journalists' free time.

It is also widespread practice in high schools to promote school journals, usually a once-a-year publication, due to difficulties in funding. Since it is not allowed to use the Ministry of Education's public budget, expenses must be covered by external sponsorship or through extra-budget funds such as mark-ups on the students' photocopies, or duplicates of the student absence justification booklets.

In this scenario, the VideoJournal project represents a different approach giving opportunities to lower production expenses for schools while promoting students' engagement as authors. The project was launched in May 2007, initially as an extension of the AffarItalianiNEWS Laboratory handout developed within the Master course on Interactive Digital Television, www.masteriDTV.unimi.it, in the Spring of 2006. It introduced an innovative newspaper format based on producing articles' content as a video itself (instead of the traditional write-up) to be read on a mobile telephone screen such as a NOKIA N70 mobile phone set just available at the time, instead of reading the same format of traditional newspaper on a PC screen, as in the online internet journal AFFARI ITALIANI, www.affaritaliani.it, (Fig.1).

Fig 1 Home Page of www.AFFARITALIANI.it

The adopted viewing technology is based on the TV over the Internet model, a widespread platform that's used in Italy for example by several banks and by the Chamber of Commerce of Milan in the information service www.impresalive.tv provided as a WEB TV, or by the private TV ventures like Televisionet, www.televisionet.tv (Fig.2), founded in Spring 2006, where many of our students currently cooperate.

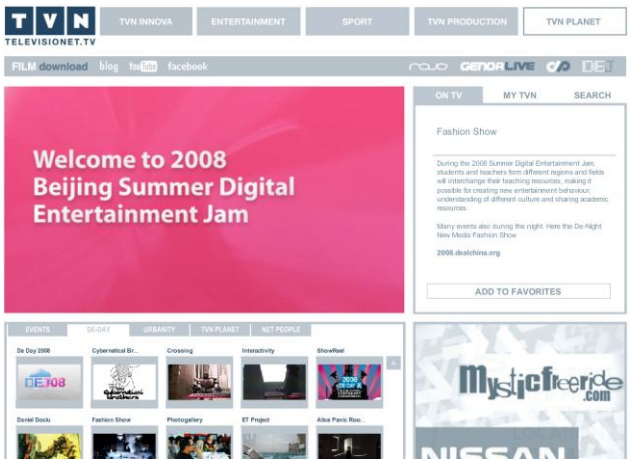


Fig.2 Home Page of www.TELEVISIONET.TV

The interactive VideoJournal consists mainly of a fusion of all the above techniques, while reversing the Giornale in Classe roles played by the students who act as:

- a journalist, when they conceive information documented by filming with a video camera
- a video and audio IT-professional, during the editing of the videos with programs like Final Cut and Adobe Premiere
- an editorial professional, by adding photos and other audio documentation.

Students, in fact, manage the video articles by themselves, completing it with the Author Name – Title – Subtitle and Flyleaf to be added (as practiced by professional editorial departments) on the VideoJournal homepage; with quite important immediate outcomes:

- authors are the students themselves,
- strong potential to promote fresh topics suited for the interests of young people
- de facto promotion of modern digital communication practices over a large base of users
- authors' cooperation with the Director of each school's VideoJournal and their editorial staff, where students' parents are also welcome
- there are no column or page limits, space is digitally unlimited
- availability is immediate, physical access depends on mobile telephone contracts and home internet connectivity to providers
- there is little power consumption, no paper waste, ink, or need for vehicles in distribution (an ecology plus)
- no need for significant budget, costs are minimal while the students' intellectual skills become the asset value
- coverage has no geographical boundary, being limited with the users' ability to understand the language that's spoken in the audio, while videos have a potential for being widely understood.

3. VGG: THE INTERACTIVE VIDEOGIORNALE GIOVANILE

The VGG (Video Giornale Giovanile in Italian) focuses primarily on the Internet, allowing it to exploit all available communication possibilities on the World Wide Web which allow immediate one-to-one and many-to-many interaction.

In fact, alongside with the single photographic image selected to appear on the VGG homepage for each video article, there are also two buttons (Fig 3) that give additional possibilities:

- a letter icon, at bottom left, allows to send an e-mail to the addresses that are linked to that specific video article, thus allowing a personalized response to the reader … and further personalized continuing communication and feedback
- a printer icon allows to print a one-page summary of the topic which the author intended to cover with that video article.

Fig.3 Interaction and Printing features.

4. CONTENT MANAGEMENT AND METADATA

To allow easy, quick and competencies-independent upgrading of the VGG from a single reading event to a repository of information , the Content Management is handled by extracting metadata from the Author Name – Title – Subtitle and Flyleaf attributes of each video article. This process has been designed following an approach similar to metadata and tags in Learning Objects methodology into an E-Learning Management System [5]. This feature allows to retrieve topics over a growing number of different subjects, as the VideoJourrnal practice will become more widespread.

5. UNIMITUBE PLATFORM

The Jurisprudence Video Journal is hosted on the UnimiTube more general website.

The UnimiTube http://video.dsi.unimi.it is a portal free accessible that allows to share multimedia video and audio files, like on YouTube but specifically targeted to University students.

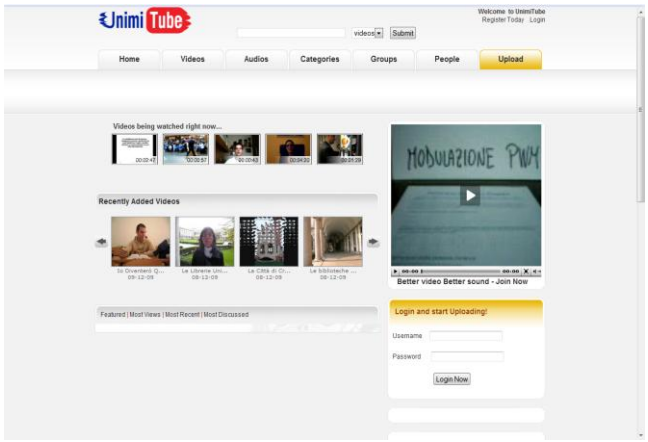


Fig 4 http://video.dsi.unimi.it

Technically speaking, the project has been developed using the free Media Sharing CMS (Content Management System)

PHPmotion (http://www.phpmotion.com) representing the kernel of the website.

The platform has been realized with a Linux (Ubuntu) distribution for server "LAMP" (Linux – Apache – MySqll – Php) providing apache2 web server, PHP 5 (with CLI support), MySQL database server 5, LAME MP3 encoder, Libogg and Libvorbis, Mencoder and Mplayer, FFMpeg, GD Library, CGI-BIN, w32codecs.

The website is divided into two main areas: the public one (at disposition of each user, bot for registered ones and for not registered) and the administrative section.

In the public area, the user, once registered, has the possibility to directly upload videos from own computer. The uploaded video can be of different formats (mpg, avi, divx and many others) after converted in flash video (flv estension). This choice allows user to see all video with the unique requirement of the installation of the flash player software, free and downloadable for all the Operative Systems. The user can modify the title, the description and the tags of the video. Moreover he/she can delete own video, make them private or public, enable or disable comments and embedding of the video on other websites. The user can also add video to own favourites, change the information related to the user profile and link to other users, creating a thematic social network.

The UnimiTube's Administrators can see video, delete them, make them "featured" and approve them, control statistics concerning videos, users, groups and comments. An Administrator can also exclude a user who doesn't follow the ethic and privacy rules and guidelines.

6. THE EXPERIMENTS WITH STUDENTS

To gain experience with students not specifically involved in the Computer Science Curriculum at the University, we report here on: a classwork assignment for senior-year students at the Law School of the University of Milan (VG Servizi Giuridici), an amusement application typical in Italian (Pesce d'Aprile Jokes) as April Fools' Day is celebrated in North America, UK, and other countries, a learning community of high school and university students (VG Learning).

VG Servizi Giuridici. The aim of the assignment was to teach video production and post-production techniques to students who had no previous educational background or experience in this field, therefore the fundamentals of shooting video films (shots, lighting, framing, etc.) and editing them through the use of computer programs such as Final Cut and Adobe Premiere were taught before launching the project. The outcome is partially available at http://video.dsi.unimi.it, while at http://vgg.dsi.unimi.it/servizigiuridici will appear the student outcome submitted on Wednesday 14 January 2009.

No restrictions were imposed regarding the subject of the videos to be made; in fact the students, some of them shy, were encouraged to choose the subjects of their personal interests. The liberty of choice regarding the subject of the videos yielded extremely positive results, as most of the students produced films about their hobbies and favorite hobbies; such as volleyball, cooking, astronomy, poker, bicycles, travel, etc.

The students were later instructed on how to upload their videos to the VGG website and edit their stories with a professional news format. These stories were later made available for their family, friends, schoolmates and anyone else surfing the internet to see them. The students appreciated the possibility of describing their personal interests and they received high grades as a reward for their good work.

Pesce d'Aprile Jokes. A similar approach has been taken with the freshmen participating in the Fundamentals of Digital Communication class of the B.A. program in Digital Communication at the School of Mathematics, Physics, and Natural Sciences in the Second Semester of the 2007/08 Academic Year. In this experiment in order to make the assignment more entertaining, it was proposed with the theme of Pesce d'Aprile (April Fool's Day). The students were asked to create Pesce d'Aprile jokes, and then shoot videos of them using cell phone cameras. The videos were later edited with computer programs for post-production, and uploaded to our video-hosting system. In order to manage the videos, a video-hosting platform with intermediate characteristics (between amateur and professional levels) and automated systems of audio/video archiving, conversion and publication was used. Once installed, this system guarantees an entirely personalized service without the necessity of paying any license fees. The Pesce d'Aprile videos created by the students can be viewed at the link http://vgg.dsi.unimi.it/pescediaprile/.

VG Learning. The following exercise provided the creation (by the students) of supportive videos for the didactics within classrooms in High Schools. Such videos were realized in collaboration with the teachers and students of the fourth and fifth years of the related educational institutions. They were created as auxiliary support to the students facing difficulties as indicated by poor grades, focusing on certain topics of particular significance for the fields related to Computer Science (mainly, the courses of Italian Language, Mathematics and Physics.)

In this phase, certain advantages related to the launch of a process of collaborative learning have thus emerged: On one side, the students of the course of Interazione Uomo-Macchina have put to practice the notions which were learned during the lessons; applying them on themes related to these studies. On the other hand, the students have found a support for their studies, particularly in line with their own needs of learning (especially for university students coming from the same secondary institutions,) experimenting in first person the qualities and the type of competencies which are acquirable at university level.

A further relapse has been concretized in the form of apprenticeships offered by the Department of Information Sciences for students who graduated from high schools. A period of time has been dedicated to collecting information regarding the necessities expressed by the teachers of the high schools. Such inquiries have been undertaken by the former students of the course of Interazione Uomo-Macchina through interviews realized at several institutions and have enabled the possibility of collecting the related data of around fifteen high schools within the urban area of Milan.

In a second phase, the groups of students will be assigned spaces on a suitable installation of the VideoJournal, envisioned for

supporting the didactic materials of diverse institutions. Each group will try to identify, based on the collected information related to the necessities, the areas in which specific courses of recovery or support for the study of one or more classes of the schools should be activated. Within such areas, didactic video materials projected with the collaboration of university tutors and high school teachers will be inserted. Complimentary activities will be related to the management of the spaces of interaction, with the participation of diverse subjects involved in the project: the university students will undertake the duties of tutor and moderator, by utilizing the instruments of collaborative learning that are made available.

The current phase of the project is characterized by the possibility to create video resources by/for teachers and students in both universities and secondary schools, either for the elaboration of contents and learning paths or for the use and sharing of materials. The first phase of the experimentation – consisting of the cooperation between the students of the University course in "Fondamenti di Comunicazione Digitale" (Fundamentals of Digital Communication) and the respective previous secondary schools (headmasters, teachers in scientific courses, students of the final year of high school) – allowed the design of a map of matters and common contents for experiences and materials. As such, tutors and students managed to create video on basic mathematics, physics and information technologies, with the active cooperation of a few university teachers.

It is obvious that a clear obstacle is the little motivation of students' and in this case we think that it could be useful to study an incentive system for the elaboration of contents in the secondary schools, both for the students and for the teachers or coordinators. There can be many advantages to be drawed, such as:

solving some problems in the relationship between the university-level didactics and those of the secondary schools (a better comprehension of respective needs in terms of arguments, competencies, levels of knowledge, and the style of teaching contents);

contributing to the improvement of the base of competencies in scientific matters;

diffusion of technical competencies and knowledge presentation thanks to the disposability of instruments (VideoJournal – video hostinmg FIG.) to create and manage video by different users (University and Secondary School students and teachers, tutors, etc.);

improvement in mapping matters and specific arguments shared at different levels by secondary schools and universities (such as Mathematics, Physics, Information Technologies, etc.);

more information regarding the University's didactic offerings, thanks to a better visibility of the specific contents of the University Degree courses, often not well-presented to the scholastic world by the usual channels, for the choice of academic studies.

Schools	study subjects	students participanting
Itis L. da Vinci	3bGeografia economica	3
Dudovich www.dudovich.org	inglese	5
I.P.C. Alberghiero "G. Falcone"	Diritto	3
I.T.C Piero della Francesca	matematica	3
I.T.C. A. De Simoni	informatica	2
I.T.C.S E. Rotterdam	Matematica	3
I.T.C.T. Ottaviano Fabrizio Mossotti	inglese	2
I.T.I.S S. Lirelli	Teoria di matematica, Matematica	3
I.T.I.S. Pino Hensemberger	basi di dati(informatica)	2
I.T.S.O.S. Marie Curie	informatica, italiano	2
IS Emilio Alessandrini	Analisi e Metodi Numerici, Matematica	7
isalgrange1@virgilio.it	informatica	
ISS Fabio Besta	Matematica	2
Istituto Gonzaga	matematica-fisica 4° ginn	3
Istituto Superiore Ettore Majorana	Filosofia , Matematica	5
Istituto tecnico commerciale Minniti	Matematica, Italiano	4
Istituto V.Gambara	Italiano	3
IT Marco Polo	Inglese	2
ITC S G. Zappa www.mczappa.it	informatica, economia aziendale, matemati	6
ITIS A. Volta	informatica 3°	4
Itis A.Einstein www.itiseinstein.it	Sistemi	3
ITIS Benedetto Castelli	Fisica	3
ITIS E. Conti	Elettronica	2
ITIS G. Galilei	Italiano	2
ITIS L. Galvani	Matematica 3A	4
ITIS Lagrange	matematica	2
ITT P.P. Pasolini	Matematica	3
L.S. Geymonat	Filosofia	3
Liceo Artistico Caravaggio	Matematica	3
Liceo Gian Battista Vico	fisica 3 A	1
Liceo Giulio Natta	Matematica - logaritmi, esponenziali	6
Liceo Linguistico G. Renier	Tedesco	3
Liceo Scientifico A. Banfi	Fisica, Matematica	5
Liceo Scientifico G. Galilei	Fisica, Matematica	5
Liceo Scientifico G. Marconi	Filosofia, Matematica, Latino	3
Liceo Scientifico L. Mascheroni	Fisica	4
liceo scientifico mariano IV d'arborea	Fisica 4-5	3
Liceo Scientifico P. Levi	Matematica	3
Liceo Scientifico S. Allende	Matematica per le classi Quinte, Fisica	4
Liceo scientifico U. Morin	Fisica,Matematica	3
Liceo scientifico Vittorio Colonna Roma	Matematica	3
Ls Giordano Bruno	matematica	3
Pino Hensemberger	informatica,Elettronica,Matematica	3
Scuola Media Prati	italiano, storia,geografia	7
44	54	147

Table A Schools, study subjects, students involved

7. SIMILAN PROJECTS AND DEVELOPMENTS ALSEWHRE IN THE WORD

Similar projects have been undertaken by numerous universities throughout the world, such as the International BBS and Internet Video Conferencing between Universities (Jost, N. Nakanishi, Y. Tatsuta, L. & Gunji, 2004) project of the Dokkyo University in Japan, which aims to develop a collaborative learning project using internet video technologies. The project allowed the students of Dokkyo University to collaborate with the students of the University of Illinois in the United States through the use of a database of video files as well as live online video conferencing and (Ministry of Education Culture Sports Science and Technology 2005) e-learning applications. Similarly, the Nagaoka University of Technology in Japan developed the Webcam-Based Knowledge Management System for Special Needs Education (Muranishi, A., Shintani, K., Kono, A., Haga, H., & Kaneda, S. 2006) project in order to create a powerful knowledge management system that archives teaching strategies and solutions for dealing with student-related issues in the classroom. The project permits the sharing of information among widely dispersed schools and other educational institutions, and supports the collaborative development of effective teaching plans(Nagamori, M., Ando, M., Nagasawa, M., Songmuang, P., & Ueno, M. 2007) ."

With the support of the LCAD lab, we explored the potentiality of the video and multimedia platforms extending the experiments also to field different from the University one. For example, the potential of Digital Video Interaction in the Tradeshow business is investigated in cooperation with UCIMU, the Italian Association of Mechanical Tool Machinery. Speeches presenting International research in LAMBDA session at LAMIERA TradeShow, BolognaFiere 14-17 May 2008, are accessible at http://vgg.usr.dsi.unimi.it/lambda. (Fig. 5)

FIG5.VideogiornaleLambda

FIG.5 VGPianetaGiovani

While the Video-Journal of PIANETA GIOVANI session at BI-MU, 3-7 October 2008 at MilanoFiera, produced together with organizer Probest SpA.is available at website http://www.polomeccanica.net/polo/index.php?option=com_content&view=category&layout=blog&id=35&Itemid=46 (Fig. 5)

8. FUTURE DEVELOPMENTS

In the past, companies such as Samsung and Sony had introduced TV sets and decoders which had the ability to connect to the Internet, but both the amount and the quality of online services which these TV sets offered were very scarce, and the keyboard-like remote control devices were too complicated for most people to use this function effectively. However, thanks to new developments in the field of computer science and engineering, it is expected that by the Spring of 2009, the Internet and Digital Television will merge in a truly multi-modal and versatile way for the first time, through a new processor that's being developed specifically for this purpose: the Intel Media Processor CE 3100, which will be used in consumer electronics such as ADSL connection capable TV sets, Blu-ray and DVD players, set top boxes for satellite television and IPTV.

This new development will make it possible to extend our e-learning support services and video journals to the interactive digital television (IDTV) network through the use of DVB-J applications and Xlets, which will also provide the necessary means for the contextual convergence with other similar content on the interactive digital television network through hypertexts. At present, the hardware of most TV sets and cell phones can't read Flash files adequately, and Intel is currently working with Adobe for solving this problem. Intel aims to place a version of Flash that's specifically designed for the CE 3100 processor by the end of 2009. This development is of particular importance for our video journal applications.

Samsung plans to offer an internet capable TV set with 8 or 9 different categories of available content, and with the ability to access local websites and internet based services such as news bulletins and weather report. Panasonic, on the other hand, will launch the Panasonic VieraCast in the United States in March 2009, which will provide access to internet-based content such as YouTube and Picasa, as well as news and sports videos, and Amazon's film-on-demand service with an archive of 40,000 available movies to choose from. Panasonic has also announced the launch of a Blu-ray player with VieraCast functions. LG Electronics, on the other hand, has introduced a new generation of TV sets and Blu-ray players with the capability of accessing the films on demand that are provided by Netflix, which can also be accessed by the Xbox 360 platform in the United States. YouTube, on the other hand, has elaborated an ad-hoc service that's specifically designed for TV sets that are connected to the PlayStation 3 and the Nintendo Wii platforms. Another new platform that was introduced recently is Sezmi, a set-top-box that's dedicated to internet-based content and is a hybrid system which brings together, for the first time, the capability to use content that's based on IDTV, IPTV and video-on-demand. [6]

The year 2012 is set by the European Union as the ultimate deadline for all EU member states to "switch off" analog broadcasts and to go "all digital" in television broadcasting. This enforced obligation, though causing several serious problems and setbacks, will enable the widespread use and fruition of our e-learning and video journal content on the interactive digital television (IDTV) network. However, in order to provide digital terrestrial television broadcasts, large numbers of ground-based transmitters will have to be installed throughout the country. The system is too costly, has several technological limitations, and has proven to be a difficult option in the Italian island of Sardinia, where the system is being tested since October 2008; soon to be followed by other pilot regions in Italy such as Valle d'Aosta, Piemonte, Trentino and Campania. Numerous experts within the major Italian broadcasting networks such as Mediaset and Rai have defined the digital terrestrial television as an "obsolete, expensive and limited" technology. Things will get even more complicated when High Definition TV broadcasts will become the new standard, as the number of available channels will have to be reduced due to the large amount of capacity allocation which HD TV broadcasts require. [7]

Likewise, the development of digital TV broadcasting for cell phones and DVB-H devices in Italy, which are also compatible with our e-learning and video journal applications, are still in an

early phase. In fact, the number of available channels broadcasting for mobile phone devices have decreased from 123 in 2007 to 114 in 2008, while for DVB-H devices, a decrease from 28 to 27 channels have occurred. So far, many channels have preferred to replicate the offerings that are available in the DVB-H network also to the cell phone network and consolidate the range of available "on demand" videos, rather than experimenting with new ad-hoc channel designs that are specifically designed for cell phone services. Nokia has entered the Italian DVB-H broadcasting market with a network named "3 Italia" that, starting from June 2008, began to offer completely free access to some of its channels; with the aim of spreading the use of DVB-H devices, and as a result, to increase their popularity among advertisement companies. The fact that the EU has selected the DVB-H as an official standard will also be helpful for the fruition of this system. However, an important moment of pause will be generated by the expiration of the contracts between Mediaset (which offers the network), TIM and Vodafone Italia that are currently in effect for the broadcasting of DVB-H and cell phone channels. The price of the rent payed to Mediaset by TIM and Vodafone Italia for using the network, and the price of the content such as movies and sports videos will probably have to be changed due to increasing pressure from the Nokia – 3 Italia partnership and its free broadcasts based on a model of financing through publicity, as well as payment for premium content. Also, the diffusion of the Mobile Web services may also increase the use of videos and other user generated multimedia content on cell phones. [8]

Several statistics, released recently by Nielsen Online, will help us to define our target audience and focus on the most feasible internet capable devices with regards to our e-learning and video journal applications, for the purpose of supporting the lectures within the institutions of higher education.

According to a research that was conducted by Nielsen Online, there were 22 million active users of the Internet in Italy in December 2008, who engage in an average of 33 internet sessions per month, spend an average of 26 hours and 11 minutes for surfing the internet, and access an average of 2011 internet pages per month. From 2007 to 2008, the average time per user of surfing the internet has increased by 32%, while the average number of pages accessed per user has increased by 41%, and the average number of internet sessions per user has increased by 14%. In December 2008, 55% of the internet users were male. However, the number of female users of the internet has increased by 4% between 2007 and 2008. The number of the senior citizens above the age of 65 who are using the internet has also increased by 14%. Still according to Nielsen, there are circa 3.3 billion mobile phone users in the world. The cell phone has a far superior level of diffusion when compared to the personal computer; in 2008 an estimated 1.2 billion cell phones have been sold in the world, circa 22 million of them in Italy; which is one of the countries with the highest percentage of cell phone diffusion in the world, as 86.4% of the Italian citizens own a cell phone (while 92.1% of Italians own a TV set). The percentage of cell phone diffusion is largest among the teenagers, 97.2% of whom own a cell phone (while 87.9% of the teenagers own a TV set.) Nielsen Online has estimated that there were 50.1 million active cell phone users in Italy during the 4th quarter of 2008; which was equal to 98% of the population with an age above 15. Only 20% of the time that's used for the cell phone is being used for vocal conversations, while 72.5% of the Italian youth are using the advanced functions of mobile devices such as DVB-H.

Apart from the SMS and MMS messages, other alternative functions that are commonly used include video games on cell phones (used by 34%, or 17.1 million of the cell phone users in Italy), the fruition of video content (used by 5%, or 2.6 million of the cell phone users in Italy), and the use of Mobile Internet (used by 15%, or 7.7 million of the cell phone users in Italy.) Meanwhile, in terms of the fruition of video content, it is revealed that 33.7% of them are being downloaded or uploaded through the Mobile Internet networks. This is followed by 31.2% of the users who use payment-based DVB-H broadcasts, while 26.4% use their cell phones to watch the downloaded video content. 20.1% of the users, on the other hand, use their cell phones to watch video content that's downloaded from another device, such as PC or iPod through a "peer to peer" environment. [9]

These statistics also mean that a cell phone or DVB-H compatible fruition of our e-learning and video journal applications will have a higher reach potential among the students.

8. CONCLUSIONS

If you consider that at the beginning of the course the students did not know almost all the technologies we teached them, you will understand our pride for the good work performed. It is also important to underline the effort done by the students in managing the project, especially if we consider the large number of students participating to the course of this academic year (about 70 students).

Through this experience the students of the course of Legal Sciences have been able to apply what they learned during the lessons on the databases, on the manipulation and management of the data and on the use of web 2.0 technologies. Of course they had to face also legal arguments linked to the publication of their works and consequently they had the possibility to explore also themes linked to their studies.

In the complex, the experimentation gave really good results, involving also the performances of the students in the final examination.

9. REFERENCES

1] Epifania, F. (2007). An experience in distance learning for university and secondary school. In G. Richards (Ed.), Proceedings of World Conference on E- Learning in Corporate, Government, Healthcare, and Higher Education 2007 (pp. 6786-6791). Chesapeake, VA: AACE.

[2] Valle, G. & Epifania, F. (2008). An Interactive VideoJournal to complement Learning & Training. In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2008 (pp. 1164-1169). Chesapeake, VA: AACE. http://www.editlib.org/index.cfm?fuseaction=Reader.ViewAbstract&paper_id=28534 [3] Valle, [3] G., & Folgieri, R. (2004). Continuous Learning Relaxing Time & Location Constraints in a Regular Curricula Degree Offered by University of Milan. World Conference on Educational Multimedia, Hypermedia and

Telecommunications 2004(1), 5166-5171. [Online]. Available: http://dl.aace.org/16225

[4] Valle, G., Reggio M. (2004), L'esperienza CampusOne e la Didattica Webcentrica del Corso di Laurea in Comunicazione Digitale: didattica a distanza e controllo di gestione, in Atti del Convegno "Tossicologia dell'Ambiente", Università degli Studi di Milano

[5] Giorgio Valle, Raffaella Folgieri (2004), Standard and ontologies perspectives in building and planning learning future, EUROMEDIA Conference 2004

[6] Longo, Alessandro: "Televisore a prova di YouTube." Published on Il Sole 24 Ore, January 22, 2009. http://qui-impresa.ilsole24ore.com/innovhub/news-sole24ore/2009/01/22/9_A_PRN.php

[7] Grasso, Aldo: "Digitale terrestre, prime sconfitte." Published on Corriere della Sera, January 30, 2009. http://www.corriere.it/economia/09_gennaio_30/grasso_433b436a-eea1-11dd-ba39-00144f02aabc.shtml

[8] Perego, Alessandro: "La TV sul cellulare cerca la via per il successo." Published on Wireless 4 Innovation, February 2009. http://www.wireless4innovation.it/web/wireless

[9] "La multicanalita' lascia o raddoppia?" Published by the Politecnico di Milano – Dipartimento di Ingegneria Gestionale; Nielsen; Nielsen Online; and Connexia, January 2009.

WORKSHOP 5

Bridging among People, Places and Devices by Integrated, Ambient and Playful Social Media Approaches

Editors:

Jan Hess (University of Siegen, Germany)
Franziska Demmig (KIT Digital Inc., Germany)
Prof. Manfred Tscheligi (University of Salzburg, Austria)
Roland Westermaier (Pixelpark AG, Germany)
Prof. Volker Wulf (University of Siegen, Germany)

An Interactive Couch Table to support TV-centric Social Interactions between Households

Lin Wan
International Institute for Socio-Informatics
53111 Bonn

Marcel Tweer
International Institute for Socio-Informatics
53111 Bonn

ABSTRACT
In this paper we present our on-going work in designing a multi-touch enabled table to better support TV-centric social activities among local and remote contacts in the living room setting. A pre-study in 5 families and altogether 9 persons were conducted. The result showed that although shared TV content plays an important role in stimulating social interactions, communications about this content during TV watching happens almost only between family members who watch TV together. A paper-mockup of the user interface design was shown to the participants and was well accepted. The system is now being developed concerning the user requirements generated from the pre-study. Noteworthy research questions and methodology to solve them are presented in this work.

Categories and Subject Descriptors
H.5.2: **User Interfaces.**

General Terms
Design, Human Factors.

Keywords
HCI, multi-touch table, TV-centric social exchange.

1. INTRODUCTION
In this paper we present our on-going work in designing and developing a multi-touch couch table for the family to interact with a media center TV collaboratively in the living room. Social aspects are also brought into the concept to better support communications between families.

In the past few years several concepts for interaction with the TV were introduced in the scientific field. Among them, the remote control approaches tried interactions with the TV per paper [4], speech [2, 3] or gesture [5, 6]. Meanwhile, the user experience in practical usage of TV and internet were observed in some work [7, 8]. The technologies of TV and internet have already begun to merge, which motivates research trials to integrate internet activities in the TV [7, 9].

The integration of internet and related services to the TV brings great challenges to the remote control. Standard remote controls do not support iTV services very well, e.g. text input [12, 13]. Even new remote control approaches have problems such as too small displays [14] or no support for social interactions [13]. Researches show that the users are switching from a passive to an active role most likely in interacting with the TV rather than with other devices [15, 16].

In the research field of multi-touch enabled tables, the iTable is exploring new interaction styles, techniques, and metaphors for collaboration around tabletops [11]. Another work presented a multi-touch table for controlling the electrical devices in the living room [10].

We are inspired by the idea of turning an ordinary couch table into a multi-touch enabled I/O device for the media center TV. With this prototype the users are able to control the media center TV, but more interestingly, collaboratively consume internet services and communicate with friends or families who have the TV content as a shared context.

2. PRE-STUDY
First of first, a pre-study was carried out to gain an explicit insight in the usage of today's media consumption in order to generate the requirements for the system from the users' perspective. Another reason for the pre-study is to involve the users in the designing process from the beginning on, in the sense of interactive design and development.

2.1 Methodology
The pre-study consists of two parts. The first part is an investigation of normal living room settings. Photos of the living room from the participating households were taken. The locations of the TV, the couch, and the couch table were investigated, in order to prove the feasibility of the system in a hardware level. The second part of the pre-study is qualitative interviews with the participants.

Altogether 5 families participated in our pre-study. And among the 5 households, 9 persons participated the semi-structured interviews. Each interview took around 15 minutes (from 12 up to 18 minutes). One researcher from our group was sitting with the interviewees during each session. The interviews were audio recorded and transcribed for qualitative analysis. The participants ranged from 17 to 53 in age and consisted of 5 males and 4 females. All the participants had middle level of technology knowledge.

Figure 1: Paper-mockup of the prototype

The interview started with questions about daily TV usage and internet communications which happen during or after the TV usage. As the second part, our researcher showed the participant a paper-mockup of the prototype interface (see Figure 1), and asked for user feedback in a first glance. Last but not the least, the participant was asked about the requirements for such system and was encouraged to illustrate some interesting use cases.

2.2 Results

After analyzing the transcripts of the interviews, we listed the results of the empirical study in categories. From them we got to know some design issues and interesting use cases.

Apart from general consumption of TV content, we found interesting results about social behaviors around TV consumption. TV watching is still a significant social event for households. All participants said that they were mostly watching TV together with family members, which approved our idea of supporting common operations around the table. The communication intensity between TV watchers in the same place depends on the genre of the watched TV program. The participants mentioned to be talking to each other more when they watch documentation, news or sports. Similar results could also be found in the work of Geerts and etc. [1]. As a further result, the participants don't communicate with distributed friends about the TV content during the TV watching. Instead, they choose to talk about the TV content face-to-face on the next day. The shared TV content is proved to be stimulating communication between participants, but this communication is mostly limited between family members who sit together in front of the TV. That's why we have the motivation of integrating some social aspects into the prototype, to better support real-time communication about the shared TV content between families.

The user feedback of the paper mockup for the system was quite positive. The most appreciated aspect of the approach is the parallel usage of the TV and social interaction. Concerning the operation of the multi-touch table, the users expressed the will of using the fingers rather than pens. Some concerns for this approach were also mentioned. Some said that the table could distract them from watching the TV, and some said they would be too lazy to operate the table for switching channels. These issues are good to be known from this early phase and should be minimized with proper design.

3. WORK IN PROGRESS

The technical goal of this work is twofold. Firstly, an extensible framework running on the multi-touch table should be developed, which offers the APIs for further plugin development. As the next step, several plugins (or we call "widgets") should be developed according to the user requirements generated from the pre-study. In the sense of UI customization, the user can drag and drop widgets on to the table, arrange their positions, or hide them for space. Figure 2 shows a screenshot of the application in early phase which should be shown later on the multi-touch table.

Figure 2: User Interface in Early Phase

According to the user requests, a remote control widget should be provided so that the user can use the table as a general remote control for the media center system. Additionally, TV-centric social widgets should be developed. An interesting use case developed by the users during the pre-study is: making a screenshot of current TV content, drawing some annotation on the table per hand, and sending it to friends in community-sites or to other families.

The concept turns the table in the living room into a public display for family collaboration, and also connects families who have these tables in their living rooms. The scenario brings interesting research questions to us. What are the potential use cases of family collaboration on this kind of public display in front of TV in a living room setting? What are the differences between a public secondary display (our work) and separate secondary displays (work of Caesar and etc. [7])? How should we handle the privacy for individuals in social interactions between families?

The methodology we will be using in solving these research questions will be usability tests with real users in a controlled lab environment in our institute, which is decorated and set up as a real living room. We will also carry out semi-structured interviews after the usability test to get qualitative results from the users.

4. REFERENCES

[1] Geerts, D., Cesar, P. & Bulterman, D., The Implications of Program Genres for the Design of Social Television Systems. In Proceedings of uxTV2008. Mountain View, California, October 22-24: ACM 2008.

[2] Stifelman, L., Arons, B. & Schmandt, C., 2001. The audio notebook: paper and pen interaction with structured speech. In Proceedings of the SIGCHI conference on Human factors in computing systems. Seattle, Washington, United States: ACM, pp. 182-189.

[3] Liu, P.-Y., et. al., An Experimental Platform Based on MCE for Interactive TV. 2008, EuroITV 2008.

[4] Hess, J., Küstermann, G., Pipek, V., pRemote – a User Customizable Remote Control, Work in Progress, 2008, CHI 2008.

[5] Huang, Thomas S. and Vladimir I.Pavlovic, Hand Gesture Modeling, Analysis, and Synthesis. 1995, in International Conference on Automatic Face and Gesture Recognition.

[6] Biswas, P. and Peter Robinson, Effects of Physical Capabilities on Interaction. 2009, EuroITV 2009.

[7] Cesar, P., Dick C.A. Bulterman and Jack Jansen, Leveraging user impact: an architecture for secondary screens usage in interactive television. 2009, Springer online.

[8] Seager, W., Knoche, H. and Sasse, M.A., TV-centric triple play services: a requirements capture. 2007, EuroITV 2007.

[9] Cesar, P., Bultermann, D.C.A., and Jansen, A.J., Usages of the Secondary Screen in an Interactive Television Environment: Control, Enrich, Share,and Transfer Television Content. 2008, EuroITV 2008.

[10] CRISTAL, Media Interaction Lab FH Hagenberg, Available from: http://mi-lab.org/projects/cristal/

[11] Stanford HCI Group, Project iTable. 2007, Available from: http://hci.stanford.edu/research/itable.html

[12] Enns, N.R.N. and I.S. MacKenzie, Touchpad-based remote control devices, in CHI 98 conference summary on Human factors in computing systems. 1998, ACM Press: Los Angeles, California, United States.

[13] Bernhaupt, R., Marianna Obrist, Astrid Weiss, Elke Beck and Manfred Tscheligi, Trends in the Living Room and Beyond in EuroITV 2007, Springer Verlag: Berlin.

[14] Seager, W., Knoche, H. and Sasse, M.A., TV-centric triple play services: a requirements capture. 2007, EuroITV 2007.

[15] Diakopoulos, N., Sergio Goldenberg and Irfan Essa, Videolyzer: Quality Analysis of Online Informatrional Video for Bloggers and Journalists. 2009, CHI 2009.

[16] Cesar, P., et al., Social Television and User Interaction. 2008, ACM Computers in Entertainment, Vol.6, No.1, 2008.

Designing a Social ITV Field Trial for real customers

Abreu, Jorge	Almeida, Pedro	Figueiredo, Cátia	Pinto, Ricardo	Teixeira, Vera
University of Aveiro	University of Aveiro	University of Aveiro	University of Aveiro	University of Aveiro
3810-193 Aveiro - Portugal	3810-193 Aveiro - Portugal	3810-193 Aveiro - Portugal	3810-193 Aveiro - Portugal	3810-193 Aveiro - Portugal
+351 234 370389	+351 234 370389	+351 234 370389	+351 234 370389	+351 234 370389
jfa@ua.pt	almeida@ua.pt	cat@ua.pt	ricardopinto@ua.pt	vmmt@ua.pt

ABSTRACT
This paper describes the specification and design of a field trial for a Social TV application. The field trial was designed to be carried out over a commercial IPTV service relying on actual customers to whom the application is offered. Social TV Games along with questionnaires, diaries, logs, direct observation and interviews are among the methods adopted to gather data related with user experience and integration of social features in the current IPTV experience.

Categories and Subject Descriptors
H5.2 [**Information Interfaces and Presentation**]: User Interfaces - *evaluation/methodology, interaction styles.*

General Terms
Measurement, Performance, Human Factors.

Keywords
Social television, interactive television, awareness, instant messaging, evaluation, methodology, usability.

1. INTRODUCTION
Research and development of solutions that enable the mediation of different forms of socialization fostered by television has been a constant. However, implementing it in the field, namely on existing IPTV commercial platforms, implies sticking to real specifications, designing for large scale use and adjusting evaluation methodologies that fit the end users that will have access to the application.

In this paper the authors report on the preparation of a Field Trial (FT) for WeOnTV - a Social TV application for supporting communication around TV content. It was developed for the most popular IPTV service in Portugal – the *MEO* product. The application relies on the integration of Instant Messaging (IM) features on television allowing users to know what others are watching, to make program recommendations or chat in multiple formats [1]. A FT with real customers is being carried out allowing close to 100 final users to try the application in their regular IPTV service. It is expected that this evaluation scenario will be more natural and rich when compared to the usage of non-regular equipments, users or contexts.

2. THE WeOnTV APPLICATION
The application runs on a regular IPTV set-top box (STB) with MediaRoom™ middleware and the corresponding browser.

A significant technical effort was put into the development of the application to overcome the hardware limitations of the STB, specifically when browser based applications and thousands of users are at stake [1]. The WeOnTV includes a set of social features like: i) flexible offer of text communication: free text input (through triple-tap insertion mode) and personalised messages (through a web application the user is able to customize the content of these messages and organize them by categories); ii) integration with public IM services (e.g. MSN and Gtalk), enabling access, in TV, to the user's IM buddy list and extending the information of presence and channel awareness to PC/mobile clients; iii) tab management of simultaneous IM sessions; iv) information on what is being watched by each buddy (privacy is also guaranteed) and; v) TV program recommendation.

3. THE EVALUATION PROCESS
3.1 Goals & Challenges
The main goal of the WeOnTV FT is related with the *in situ* evaluation of the usability and user experience of the referred application. Despite some important FT like the in-depth field study of STV3 [3], the evaluation of Social TV applications over commercial and running IPTV services is still far from a stabilized approach. Therefore it is worth to discuss in this paper some ideas about the planned methodological set up. This study will also report on the viability of other (social) features to be included in a future version of the application (e.g. cross-platform content recommendation).

3.2 Methodology
The methodology of the WeOnTV Field Trial relies on 3 interrelated phases (see figure 1).

The first phase is aimed at the characterization of the evaluators and their association in groups (aiming to evolve into communities). This is a critical step since for the IPTV operator it is important to have a friendly and non intrusive approach to their customers. A brief phone contact, done by the operator aims to identify customers willing to participate in the FT. They will also be asked to point out friends (*MEO* customers) that may be available to participate in the FT. It is expected that this Friend-of-a-Friend approach will help the evaluators build clusters of participants (small communities of friends). After the definition of a large group of participants the research team will apply a questionnaire in order to characterize them as to digital literacy and level of technology adoption.

The second and central phase relies on an articulation of

"traditional" methods and creative and playful techniques (based on gamming challenges).

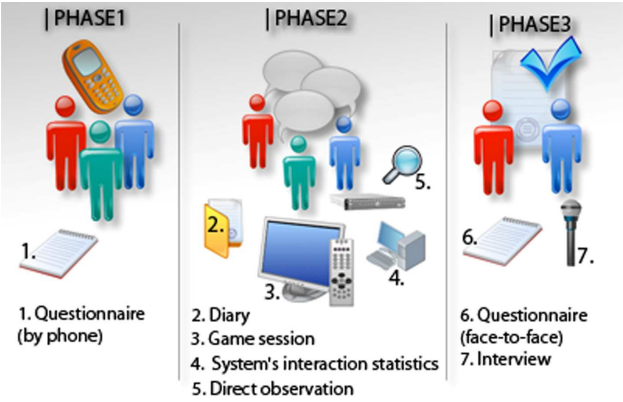


Figure 1. WeOnTV Field Trial phases

The set of "traditional" methods is based on: a personalized diary (the users are asked, on a daily basis, to answer a few questions and to classify, using emoticons in a *"fitting diary template"* [2], topics related with their social and TV experience using the application); direct observation; system logs and a final questionnaire and interview. Considering the system logs, the interaction in the client application will be tracked and a set of graphics will allow real time monitoring of the users' activities.

Considering that the field trial will run in an "uncontrolled" environment (the households), some quizzes related with TV live events are planned to promote the social use of the application. These games are aimed to increase social interaction around the TV content as it is believed that *"participants can be involved more actively in the whole flow (...) and enjoy a more comfortable atmosphere"* [4]. For this, the selected participants must track a buddy (common to all evaluators) that acts as a game presenter; search and invite other friends to establish a team; find other game instructions and ratings on a dedicated blog; and interact using the available features (namely communication via fixed and free messages and TV program recommendations) to answer the questions and challenges promoted by the game presenter.

Finally, the third and last phase is aimed at collecting final data through a questionnaire and face-to-face interviews.

3.3 End (*MEO*) users as participants

The evaluation process will be applied at two different moments to two different types of users. The first moment acts simultaneously as a pilot evaluation with the participation of 40 Portugal Telecom collaborators. This group of *MEO* pilot evaluators has high digital literacy and experience in testing new *MEO* applications. Most of the participants know each other, a fact that probably will act as facilitator as far as community dynamics is concerned.

After this first evaluation some technical fixes to the application or adjustments to the data collection instruments may be performed. If necessary these can be carried out at any time since the application relies on a server oriented model.

The second evaluation moment is targeted at regular *MEO* customers from the Aveiro region. The geographic factor is only considered to ease the researchers' task of delivering the diaries and perform direct observations at the home of some participants.

It is expected that a group of about 60 *MEO* customers will be enrolled in this evaluation. No technical modification by the *MEO* staff will be needed in households, since it is only necessary to remotely activate the access to the WeOnTV application in their IPTV set-top boxes.

4. PRELIMINARY CONCLUSIONS

The WeOnTV evaluation was designed to fit both academic and commercial purposes. Therefore, an *in situ* evaluation was designed. This has brought several challenges as to technical developments (the need to have a robust, crash free application in the regular and commercial IPTV infrastructure) and adapted methodological design. The evaluators need to have limited interaction with the participants in their home environment but a high level of engagement is targeted to fully experience the application. Therefore some TV content related games were developed. The researchers expect that this approach together with more "traditional" evaluation methods will allow getting closer to reality results as compared with in lab evaluations, keeping the rigorousness of data collection.

A future publication will present the results gathered using this methodology and will report on the evaluation and validation of the proposed methods. Important results concerning the users' expectations and actual use of a Social TV application concerning Portuguese viewers are also expected.

5. ACKNOWLEDGMENTS

Our thanks to the other WeOnTV research fellows M. Artur and V. Nobre. Thanks also to the PTIn and labs.sapo/UA teams.

6. REFERENCES

[1] Abreu, J. F., Almeida, P., Pinto, R., and Nobre, V. 2009. Implementation of social features over regular IPTV stb. In Proc. of the Seventh European Conference on European interactive Television Conference (Leuven, Belgium, June 03 - 05, 2009). EuroITV '09. ACM, New York, NY, 29-32.

[2] Hess, J. and Wulf, V. 2009. Explore Social Behaviour around Rich-Media: A Structured Diary Study, In Proc. of the Seventh European Conference on European interactive Television Conference (Leuven, Belgium, June 03 - 05, 2009). EuroITV '09. ACM, New York, NY, 215-218.

[3] Huang, E. M., Harboe, G., Tullio, J., Novak, A., Massey, N., Metcalf, C. J., and Romano, G. 2009. Of social television comes home: a field study of communication choices and practices in TV-based text and voice chat. In Proc. of the 27th international Conf. on Human Factors in Computing Systems (Boston, MA, USA, April 04 - 09, 2009). CHI '09. ACM, New York, NY, 585-594

[4] Obrist, M., Moser, C., Alliez, D., Holocher, T., and Tscheligi, M. 2009. Connecting TV & PC: an in-situ field evaluation of an unified electronic program guide concept. In Proc. of the Seventh European Conference on European interactive Television Conference (Leuven, Belgium, June 03 - 05, 2009). EuroITV '09. ACM, New York, NY, 91-100.

Designing a large social display for an old people's home

Claudia Müller
Graduate School Locating Media
University of Siegen
claudia.mueller@uni-siegen.de

Cornelius Neufeldt
Graduate School Locating Media
University of Siegen
cornelius.neufeldt@uni-siegen.de

Lars Schröer
Graduate School Locating Media
University of Siegen
lars.schroeer@uni-siegen.de

ABSTRACT

This paper reports on a pre-study which has been accomplished in the context of a design project for a large-screen display in a German retirement home for the elderly. The vision for the large social display to be built is to foster social interaction among the people, to serve as a collaboratively used information tool and also to provide entertainment, such as playing games and watching TV together. We provide some reflections on implications for the introduction, design and adoption of new social media in an old-people's home. We discovered issues important for the following project steps such as figuring out a way how to communicate and place the media in the course of the peoples' day structure or also that self-presentation is an important issue still in higher ages.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems - Evaluation/methodology; H.5.2 [Information Interfaces and Presentation]: User Interfaces - Prototyping, User-centered design; K.4.2 [Computers and Society]: Social Issues - Assistive technologies for persons with disabilities;

General Terms

Design, Human Factors, Theory

Keywords

Advanced age, social display, social media, multimedia, media acceptance, social interaction, old people's home

1. INTRODUCTION

This paper reports on a pre-study which has been accomplished in the context of a design project for a large-screen display in a German old-people's home. The vision for the large social display to be built is to foster social interaction among the people, to serve as a collaboratively used information tool and also to provide entertainment, such as playing games, watching TV together etc. We accomplished on-site interviews in order to gain an understanding of how every-day life is being conducted by the residents and in order to elicit aspects in social activities and concurrent media use for gaining ideas for framing future applications. As the residents do not have internet access and thus no knowledge on how the internet could be a multifaceted resource for them, we arranged in the course of a student's seminar an "internet day".

In the following chapter we give an overview on related work. The subsequent chapters give a short introduction of the setting – the old-people's home – and the research methods. In the findings chapter we first report on the acceptance or denial of the different offers at the "internet day". The following chapters frame these experiences with results from data gathered in the interviews before and after. We then provide some preliminary reflections on implications for the introduction, design and adoption of new social media in an old-people's home. We discovered issues as important for the following project steps such as the way how to communicate and place the media in the course of the peoples' day structure or that self-presentation is an important issue still in higher ages.

2. RELATED WORK

The ageing of the society has meanwhile stimulated many research projects which center on designing ICT and new media for the elderly for improving health and social support and staying socially connected. Especially social media have the potential to develop new ways of communication, information, and social interaction for elderly people who suffer from physical, psychical and social health problems. However, many elderly people are still left out from social media research approaches [1]. The main challenges at the interface of social media and elderly citizens is the need to acknowledge the huge diversity of the senior population in terms of life styles and living contexts which bear different needs and demands. The corresponding varieties of interactive services and resources which may be provided by social media call for user-oriented research approaches and methods which have led to important insights concerning sociability and user communities in different research areas, such as from iTV and CHI research, e.g. [2], [3]. However, whereas research in the field of ambient assisted living aims at supporting an independent life at home as long as possible (and which certainly is an important direction), other spatial living contexts, such as conducting the later life in an old people's home has been underrepresented in social media research up to now. The same is true for research on large displays, which are contextualized in public and semi-public spheres and also in the home, but which have not been thought of as interaction media for elder communities or old people's homes up to now, e.g. [4], [5]. Another branch of research our project touches on is work on web accessibility for elderly persons. There are various attempts to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
EuroITV2010, June 9th-11th 2010, Tampere, FINLAND.
Copyright 2010 ACM 978-1-60558-831-5 /10/06…$10.00.

make the internet more usable for senior citizens, such as [6]. Again, the certain living conditions of seniors of advanced age need certain attention in access of leveraging physical problems such as fine motor and visual skills. As to design methodology there is a large corpus of action research and participative methods in media design for the elderly. Though the old people's home deserves modified methods - at least in Germany - as being commonly seen as far away from interests in media and internet usage. As we believed the internet also proves as a valuable resource for social media applications for seniors of advanced age, but did not know how and in which form, we introduced "the internet" there as a kind of an open qualitative experiment. In the following some preliminary experiences will be reported and discussed as ideas and early implications for the communication, design and adoption of the social display and according meaningful applications.

3. THE OLD-PEOPLE'S HOME

The study and the "internet day" event took place in an old people's home in a small German city. The home houses about 90 elderly who mostly suffer from physical or mental diseases and who are mainly of advanced age (80+). As many people only move into the house when living an independent life at home is no more possible they often enter the house with severe problems. Due to their higher age and according health problems which e.g. cause decreasing mobility or social isolation often over years in their homes, building new relationships in the new circumstance is hard or of lower interest to them and also interest and motivation in actively taking part in social life often declines.

The house offers many regular group activities to their inhabitants.

The seniors did not use the internet before, the main media they consume in different extents are television, radio and newspapers. For two years play sessions with a *Nintendo Wii* have been regularly offered by the social workers (every 1-3 month) and are being visited by some inhabitants.

4. METHODS

We selected an empirical approach for the project in the old people's home in order to get data grounded in real practices and implications for design. We started with two days of observations in the old people's home. Our focus was on the everyday life and behavior of the elderly before introducing new ideas or technologies. After this period of time we discussed the results in workshops and found the findings by conducting eight semi-structured interviews with the elderly, querying them about their life and their relation towards new technologies.

After those interviews we planned the "internet day" based on the outcome in order to show the elderly the possibilities new technologies can offer them. We decided on this kind of qualitative experiment because we noticed that after the interviews and discussions we still were confronted with a lack of understanding of the new possibilities. During the internet day we presented the different online services at five stations with a large monitor (24"), four with computers, and one with the *Wii* console. We developed concepts for the stations defining contents to show argumentations and ways of explaining the technologies towards the elderly. The content services were *Skype*, *Facebook*, media libraries, and *Google Earth*. The stations were situated in the lobby where the elderly usually have to pass by when leaving the building or going to eat a meal. The lobby also provides seating-accommodation which enables a chat with persons passing by. The event was documented by video, audio, photos and notes that were taken during the event.

After the event we did 15 interviews with the elderly, caregiving personal and the management of the old people's home to evaluate the effects of the event. The interviews were transcribed and analyzed with an adapted method of the contextual design approach by three project workers.

5. SCHEDULE – THE INTERNET DAY

The event was planned to take place on a Saturday morning where no other regular offers are usually planned for the elderly. It was promoted as a special event in the old people's home and in the village where the home is located, by passing out flyers to the elderly and by hanging posters in the building. The project group consisted of a student group of 8 persons and of three lecturers. The event was planned to start at 10 a.m., but when we started to prepare the desks and switch on the computers at 9 a.m. there were already many elderly coming down to see what was happening here. During the event the elderly were mostly brought to the stations by the people from the social services and the manager of the old people's home. They also had actively promoted the event towards them beforehand. The elderly who did not take notice of the posters noticed that there was something going on by other signs e.g. more cars in the parking lot and more younger people in the home, so they too became interested in the event. During the entire event we had many elderly actively participating. Regarding the handling of the mouse and keyboards, all elderly people wanted the students to operate the devices although they have been repeatedly asked to try it out themselves.

6. EMPIRICAL RESULTS

During the study we found different aspects which are important to understand the life of the elderly and their relationships toward technology.

6.1 Acceptance of the different internet stations

The acceptance and the way of taking part in the different stations were very different. One effect in the participation at the stations was that a few dominant elderly took the attention of the students leading the station and the elderly who were not dominant had only few chances to get in touch with the new technologies. We have to see this observation in correlation to many elderly who were not willing to take part actively and were only watching what others did. The students tried to involve as many of the elderly sitting at their stations as possible, which was not always possible because of the masses of elderly who were curious about the event. The curiosity was towards different aspects which were the event itself, the young people being interested in interacting with the elderly, the technology we were presenting and the contents shown on the computers. We will show some examples for different acceptance of the stations in the following paragraph.

At the station where we presented internet video chat with *Skype*, we were confronted with a lack of acceptance of the new technology because of problems in the setting. We were offering

Skype in a setting were both PCs were in the same room and near to each other, and our idea was to show how easy you can video chat to each other. We were aware of the fact, that the setting was very artificial and explained this to the elderly. But the elderly could not find a sense in video chat because the person they were talking to was sitting in the same room, so they did not see any use in the new technology. Aside from this problem in understanding our intention, we could observe that many elderly were fascinated when they found the person they talked to in the computer in the room.

At another station we presented the online media library of a German television channel. Our intention was to show news and content from the neighborhood of the old people's home as we found out in prior interviews that the elderly were mostly interested in those news stories. We also had evidence from other studies that some content in the library was of special interest for people above 50 years of age. At the internet day we noticed that the content was not as accepted as planned. We found two reasons explaining this fact: First was that the group of 50+ is a large group of individuals with different interests. So, for the future, we have to define smaller groups when we try to find interesting content. The other fact was that the elderly have quite a bit of on their hands time to consume TV content and consequently have no need for getting content they missed. This can be seen at the example of clips from the Olympic Games which were running during the study. The elderly explained to us that they have enough time to keep awake to see the Olympic games coming at night in the VT, so they don`t need the recorded clips from the digital library.

At the *Facebook* station we could see many interested elderly searching for relatives and writing comments to them. They also had fun seeing the account of the manager of the home and leaving a message on his wall. We presented Facebook as the "yearbook of today" and the seniors had no problems understanding the offer. The elderly knew yearbooks very well and were interested in the content.

The most intensive use could be observed at the stations where we switched to use *Google Earth* and web pages from different cities to give the elderly the possibilities to see pictures from their cities of birth again. We had no plan to show this content but the social caregivers of the home planned and announced this offer to the elderly, which shows that a deep understanding of the everyday life, needs and the background and interests of the persons is essential for a good development of offers. As a positive example for the success of this station we could name a 98-year-old person who was on the internet the first time on this day and was very happy to be able to see his place of birth and the company where he worked for almost 50 years again after a long time of no contact to this area.

At the *Wii* station the students and some seniors started with *Wii Bowling*, a game often mentioned to be very well accepted and liked by elderly. However, they quickly noticed problems with handling the controller when being physically handicapped, as described in [7]. They switched to *Wii Sports Resort* game and chose the game *"Air Sport"*. Controlling a plane without time pressure and slow and small arm movements proved to be more attractive. Playing the Bowling game, in contrast, to some had caused a fear of failing due to problems with movement and fast reaction. Another reason for the higher acceptance of the flying game was that there was no real target and by this there was no winner or loser.

In the event we often noticed that at the beginning the technology was in the focus of interest by the elderly because it was new and unknown. But they quickly forgot about the technology when they came in contact with the content. So we could observe that the content is more important than the technology if we want to create new offers.

6.2 Finding a place in the every-day structure

"It was something new, but I see the inhabitants' reluctance in accepting new offers. Some do not want new activities, because they are lazy." A 60-year old lady who is very active in her life conduct said this about her co-inhabitants who are mainly much older and who could be more active from her point of view and who do not want to engage anymore in active life. Her personal attitude equals to findings from the gerontology literature: Due to health-economic reasons in Germany, there is a new trend in the last years: elder people often stay at home as long as possible and go to an old people's home only when there is no other solution for their care. That's why they often come with reduced bodily and mental capabilities and often refuse participation in activities.

However, our "internet day" aroused so much attention and had been visited by so many people of whom the social workers were totally surprised and from whom they had not expected so much interest: "we see here that we often have too little confidence in capabilities and interests of our inhabitants". There was a bias in the interpretation of that phenomenon in the social workers' and other participants' view: on one hand regular daily and weekly appointments are important for many of the inhabitants to structure their every-day life, such as fixed times for meals and activities such as a daily newspaper reading group, a gymnastic circle once a week or a film evening once a month.

On the other hand, some interview partners said that new activity offers often are accepted very well, but motivation to participate in follow-up events or regular events is then very low. The motivation for the participation in the first event to some seems to be a certain curiosity, an interest for everything that is new and which contrasts the every-day life which otherwise is very regular and unvaried.

Preferences in socializing with other inhabitants also influence the acceptance of the activity offers in the old people's home: Some people like to interact regularly with others, but some don't. Anger and dispute, as in other living contexts, also belong to the life in an old people's home. Thus, some people do not like to participate in activities when they know persons they do not like come there, too, even if they would have liked the offer. Another reason for not participating was given by an old lady who said that she would disturb the gymnastics group because she needs to go often to the restroom due to her certain uretic medication.

6.3 Looking good is still important among 80-year olds!

One elderly lady admitted that she had "*exercised theoretically*" the arm movement for the *Wii* bowling game. She had participated at a *Wii* bowling session which had been organized by the social workers before the internet day. At that time she felt embarrassed when she could not manage to press the buttons as

required. After that session she did not participate in that event again, even when the social workers invited her to. After having exercised the movement she had the feeling she could dare to try it again with us.

Another lady told us "*playing the game was so much fun because I noticed that I could do it. I first thought I was not able to do it, but then I was really surprised and happy that I made it.*" This utterance of the feeling of success is an indicator for a problem which is experienced by many elderly people with physical handicaps: they often feel uncomfortable to expose themselves in public and even in an intimate circle of other inhabitants assembled in front of the game. This feeling often is a hindrance for participation in activities in the house. From another perspective the positive feeling in the game was not only fun for her but also gave her back a little trust in her physical capabilities.

7. PRELIMINARY IMPLICATIONS FOR SOCIAL MEDIA INTRODUCTION, DESIGN AND ADOPTION

Bring the technology with you!

In participatory design projects with persons who don't know the technological possibilities there is no other way than to bring the current technology to them. We could see in our projects that the ideas in the interviews after the event were more founded and better formulated than before the event. This optimization was regarding both the technology and the content of the ideas.

Designing games for elderly

Games for elderly have to follow different rules. One the one hand they have to deal with special problems regarding mental and physical capabilities. On the other hand they should not require hectic and fast reactions from the elderly; they should be slow and fault tolerant. As we could observe that self-presentation is still important for elderly, we have to care that the use is easy enough and that everyone can handle it directly without being stigmatized. We also have seen that games such as *Wii* have the potential to evoke feelings of success and to reactivate trust in physical capabilities. A deeper examination of these factors for designing games and other social applications is needed.

Rethinking roles for handling the system

Besides the need for designing easy-to-use input devices for persons with handicaps such as fine motor skills it also became obvious that we need to rethink the user roles. As the social workers currently overtake an important role in organizing and leading social activities in old-people's homes, they should be integrated in the modeling of user roles for the prospective system, too. Additionally, this seems important due to the fact that many elderly people liked to being shown the different internet contents and programs but all refused to operate the PC themselves. Even though we attempt to provide devices which can easily be used by everyone, the social aspect of the use situation must been more strongly taken into account for that special setting.

Highlights vs. regular activities

The different approaches of dealing with "activities" and "events" in the old people's home – some like regular events, some come to "highlights" but not to follow-up events - is important information for us when thinking about the introduction of the large screen display and according applications. Technology acceptance among elderly people is an important issue and from our point of view related success or fail already relies to the way of how the display and the applications are introduced and presented according to each conduct of life of the inhabitants. Our pre-study points to the fact that we should frame the display activities not only as regular offers but also as certain "highlight events", in order to be attractive and meaningful for as many people as possible.

Search for points of contact!

The huge interest for searching the birth place and other information of their former phases of life in *Google Earth* and on other web pages indicates the importance for searching points of contact between individual life stories and potentials of new media. A motivation to go in for new media can only be tightened when points of contact in terms of identification with the topics have been found.

8. ACKNOWLEDGMENTS

Our thanks to the elderly people who shared their experiences with us and to GSS GmbH Siegen, the management and the social workers of the Haus St. Elisabeth for the excellent support during our studies.

9. REFERENCES

[1] R. Borrino, M. Furini, und M. Roccetti, "Augmenting social media accessibility," Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibililty (W4A), Madrid, Spain: ACM, 2009, S. 54-57.

[2] S.E. Lindley, R. Harper, und A. Sellen, "Designing for elders: exploring the complexity of relationships in later life," Proceedings of the 22nd British HCI Group Annual Conference on HCI 2008: People and Computers XXII: Culture, Creativity, Interaction - Volume 1, Liverpool, United Kingdom: British Computer Society, 2008, S. 77-86.

[3] M. Obrist, M. Miletich, T. Holocher, E. Beck, S. Kepplinger, P. Muzak, R. Bernhaupt, und M. Tscheligi, "Local communities and IPTV: Lessons learned in an early design and development phase," Comput. Entertain., vol. 7, 2009, S. 1-21.

[4] Churchill, E. F. ; Nelson, L. D. 2007. Interactive community bulletin boards as conversational hubs and sites for playful visual repartee. 40th Annual Hawaii International Conference on Systems Science (HICSS'07); Seventh Annual Minitrack on Persistent Conversation; 2007 January 3-6; Waikoloa, Hawaii

[5] Y. Yoon, H. Ryu, J. Lee, S. Park, S. Yoo, und S. Choi, "An Ambient Display for the Elderly," Universal Access in Human-Computer Interaction. Ambient Interaction, 2007, S. 1045-1051.

[6] V.L. Hanson, "Web access for elderly citizens," Proceedings of the 2001 EC/NSF workshop on Universal accessibility of ubiquitous computing: providing for the elderly, Portugal: ACM, 2001, S. 14-18.

[7] Neufeldt, C. 2009. Wii play with elderly people. In: C. Müller & M. Lewkowicz: Enhancing Interaction Spaces by Social Media for the Elderly. A workshop report. (International Reports on Socio-Informatics) (IRSI) Vol. 6 Iss. 3.

Constraints and simplification for a better mobile video annotation and content customization process

Robert Knauf
Chemnitz University of Technology
Strasse der Nationen 62
09111 Chemnitz – Germany
+49 371 531-35793
robert.knauf
@cs.tu-chemnitz.de

Arne Berger
Chemnitz University of Technology
Strasse der Nationen 62
09111 Chemnitz – Germany
+49 371 531-36872
arne.berger
@cs.tu-chemnitz.de

Maximilian Eibl
Chemnitz University of Technology
Strasse der Nationen 62
09111 Chemnitz – Germany
+49 371 531-31562
maximilian.eibl
@cs.tu-chemnitz.de

ABSTRACT
Users have limited will and creativity for describing, tagging, and rating audiovisual content, especially when they are consuming media in a mobile setting. The approach of our research is to reduce the burden of intellectually annotating user generated videos and to simplify content rating. Thus, more relevant results and more accuracy of fit are to be gained when potential consumers are in search of quick entertainment. We propose a video upload tool for smartphones which allows users to share captured video clips with buddies and the community. The distinctive feature is an elementary, lightweight but expressive ontology that speeds up and simplifies content tagging for the producer on the one hand, and allows very quick and well-matched retrieval of potentially interesting media items on the other hand. Furthermore, an unobtrusive but required rating request should lead to more precise estimations regarding the relevance of content.

Categories and Subject Descriptors
H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – *Evaluation/methodology*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Evaluation/methodology, Graphical user interfaces (GUI), User-centered design.*

General Terms
Design, Experimentation, Human Factors

Keywords
mobile video, mobile television, video annotation, video retrieval, video search, rating, GUI

1. INTRODUCTION
The way of selecting and consuming audio-visual content in mobile settings varies widely from "ordinary" domestic usage. Most often the users are in a hurry and have only limited time for enjoying entertainment on their mobile smartphones while on the move. Whether they are commuting or going out with friends, users usually avoid constructing complex search queries to find suitable content to watch. In the same way, they often omit rating media items after they are consumed. Besides time constraints, factors such as a lack of concentration, obtrusive user interfaces, and constrictions from inter-personal circumstances (e.g. commuting in a crowded train or quickly retrieving a funny video clip to entertain a clique), may prevent a search engine from gaining insight into the entertainment needs of the users.

An even greater amount of creativity and potentially disruptive interaction is required when describing content one has just captured in a rich and sufficient way for others to retrieve it. Superficially speaking, entering tags or keywords grants freedom of description. Despite having that freedom, people often describe being unable to think of any tags [8]. Nevertheless [5] wrote that user-based tags are usually derived from a connection between concept and content in the mind of the author. They assume that tagging supports solely the supplier of the metadata or just a few users because of ambiguities or even codes in closed user groups. As a result, a loss of meaning may occur beyond the context of a single user. Neither the community nor the recommendation systems take advantage of uncontrolled and chaotic tagging behavior when they look for relevant content afterwards. This is aggravated by the fact that a huge number of terms entered in tagging applications are misspelled, mistakenly encoded, or given in a plural or otherwise altered form or as compound words. [5] point out that a serious weak point in tagging audiovisual content is that most users do not give much thought to the way they tag resources. Changing tagging habits turns out to be a difficult task, because people tend to repeat behavior they have acquired in the past [8]. As an exception, observing community tags has an influence on an individual's tagging style.

In addition, an oversaturated vocabulary is deemed less efficient for retrieving adequate content from a huge amount of information, as observed for the social bookmark tagging service, *del.icio.us*[1] in [4]. With an increasing level of tag saturation, users must either find new words or use multiple words to describe introduced content. But [5] illustrate that most tags are generally used by one, two, or small groups of users. The number of tags that are used by the whole community is relatively small. Over time, increasing numbers of available tags diminish the completeness of search results.

Current research mainly focuses on automatic video recommendation systems to find relevant video clips according to the users viewing preferences [7] based on the recurring

[1] "*Delicious – Social Bookmarking*": http://del.icio.us/

assumption that existing tags in video repositories are quite rare [2]. Users' conscious decisions are often undermined by recommender algorithms that try to be context-sensitive, as described in [3]. Video sites like *YouTube*[2] try to recommend relevant video clips as soon as users finish watching a clip. The system does not communicate the basis upon which a recommendation has been made to the users. Consequently, the algorithm might carry the user farther away from his entertainment source.

With regard to this question, we proceed from the assumption that finding interesting media items quickly and matching the users' entertainment requirements at just the right moment is difficult for both the content provider and for the requesting consumers. The negative effect of the lack of both descriptive metadata and broad user feedback is difficulty in retrieving suitable content inside a video portal with as few clicks as possible. Most often, the seeking consumers are left having to refine their search query several times or to use non-customized item lists such as "most viewed". Most commonly, retrieving media items referring to a local geographic area, a special interest, or a social group can only be achieved by following the "channels" of certain users or user groups. Social media sites such as *Flickr*[3] enable users to keep track of their contacts' content both by harvesting activities of interconnected users and by providing interfaces that create personalized "Explore pages," offering interesting content produced by the users' contacts [6]. Feasible solutions of this kind for overcoming information overload reduce the importance of tags as a way to share multimedia content.

2. ANNOTATING AND RATING QUICKLY BUT RESTRICTIVELY

To improve the lopsided ratio of content and assigning metadata, or tags, we propose a tool for mobile phones, such as smartphones, with features as follows:

The basic functionality includes connecting to an online video platform. The participants create their own audiovisual content using the built-in cameras on their mobile devices for the purpose of sharing impressions with buddies or with the whole community. Suitable platforms include *Qik*[4] or *YouTube* to name but a few.

In addition, we focus on enriching the repository of content with descriptive metadata such as tags, categories or feedback information. Participants are to be increasingly required to make their own content more relevant in associated search or selection procedures. Concurrently, we want to increase tagging activity by providing methods of quick and easy annotation at the point of capturing content, as previously suggested by [1].

2.1 Mandatory but quick and easy tagging

Our proposed concept uses a balance between constraints on the one hand, and unobtrusiveness on the other: The video upload procedure permits transferring data to the server only if proper metadata has been provided by the user. So, capturing the video clip and entering appropriate tags should not take more than three steps of screen interaction with the purpose of limiting usage interruption. The strict limitation of user interaction is achieved as follows.

2.1.1 Limiting descriptions to cardinal questions

As suggested by [8], in persuading a community, it is helpful to specify only certain types of tags that benefit the system. Their results suggest that users would tend to follow a "pre-seeded" tag ontology.

For generating narrowed but important content descriptions, only metadata is considered that answers the questions *Who?*, *What?*, *How?*, *Where?* and *When?*

For the questions *Who?*, *Where?* and *When?* it is usually possible to retrieve the answers automatically. Users identify themselves to the video portal by providing a username and a proper (often locally stored) password. If personal identification is not wanted, one might optionally publish content anonymously or as a member of a user group. Furthermore, a mobile device with positioning capabilities (e.g. GPS, Wi-Fi triangulation) provides both location and time.

Only the questions *What?* and *How?* remain to be answered manually. For answering these questions we propose to answer *What?* with a descriptive noun, representing a category, and *How?* with a corresponding adjective. Figure 1 shows the dummy application we used for user interviews.

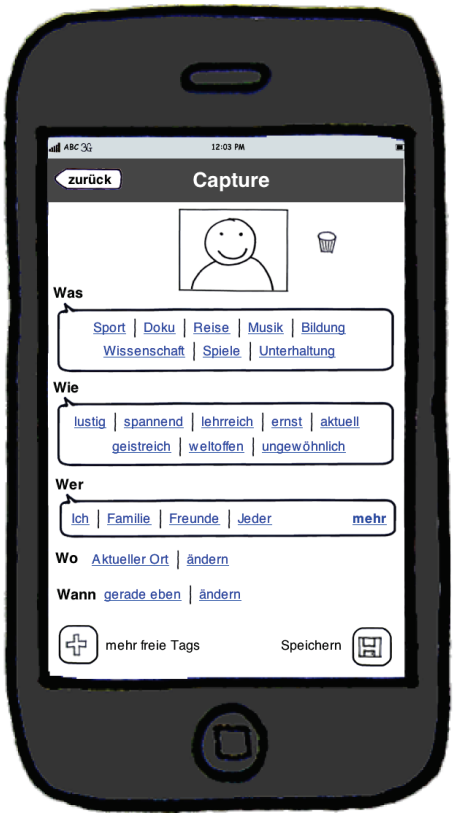


Figure 1. Dummy prototype for entering tags after video capture

[2] "*YouTube – Broadcast yourself*": http://www.youtube.com/

[3] "*Flickr – Photo Sharing*": http://www.flickr.com/

[4] "*Qik – Record and share video live from your mobile phone*": http://qik.com/

2.1.2 Limiting keyword vocabulary

As previously described, one can anticipate limitations in users' will and creativity to provide descriptive keywords when uploading content. Adverse factors such as distraction, shortage of time or unwillingness often keep them from submitting specific tags.

For these reasons, manually entering tags has been omitted, and instead, users choose appropriate keywords from a closed and narrow vocabulary. Since uploading a video clip and providing metadata should take place with no more than three steps requiring user interaction, as previously indicated, only one keyword per remaining cardinal question (*What?* and *How?*) is to be selected.

2.2 Rating before carrying on

Our second proposition for gaining relevant user feedback in an only slightly annoying manner is to reuse the principle of constraints and simplification. After watching a video clip, a consumer may only proceed to the next item after submitting a rating. It should only require one click to accomplish this. Additionally, a choice of "I don't care/know" is to be kept available.

3. REQUESTING INTERESTING CONTENT

The same policies proposed for upload functionality in 2.1.1 and 2.1.2 are to be applied to the search process as well: One can shape a simple search request with just a few steps of interaction with the graphical user interface of the software. The request is to be constructed from the terms that have already been predefined for video annotation (see 2.1.2). For this purpose, the cardinal questions described above (see 2.1.1) provide an elementary structure for creating requests to find interesting media items.

Through this process, the need for entertainment is channeled into a strongly structured query with the objective of delivering an individualized media "channel" to the user that contains all items matching the user's presumed interest. Moreover, users will be able to customize the results by adding restrictions or defining sort criteria such as geographic location, topicality or content provided by certain community members.

Both search queries and result order are to be generated by using the same metadata set. The graphical user interface of our proposed tool shall provide functionality to interrelate primary search terms from one class (e.g. *What?* or *How?*) with limiting parameters derived from one or more other classes (e.g. *Who?*, *Where?*, *When?* or popularity, see Table 1 for examples). Figure 2 again shows the corresponding dummy application used for user interviews.

Table 1. Examples of using cardinal questions to limit or sort search results

Question	Default terms
Who?	single person, friend, family member, anybody
Where?	current location, home, known POI, map position
When?	last *n* hours, days, weeks, etc., time span

Assuming that a user's interests remain more or less stable, an interesting potential use is to follow events related to a certain person, location or time span.

4. RESEARCH QUESTIONS

Our approach is to be tested first by paper-prototyping and by setting up an initial set of keywords. These frontend prototypes will be tested by potential users (see Figures 1 and 2). In the subsequent focus group interviews, we want to figure out if the participants accept the idea or not. For this purpose we will draw a comparison between our concept and the ordinary "freehand tagging."

We anticipate acceptance if the advantages of the simplification methods prevail over negative impressions resulting from the application of constraints.

With the knowledge gained from the interviews we will implement the tool on a mobile phone software platform coupled with a portal server system. In subsequent user studies we want to answer the following questions: Is the association between given keywords and actual content correct? Is the set of abstract keywords sufficient? Is the concept of constrained rating providing better and more realistic scores or is it necessary to add editorial scoring?

Finally we want to find out if and how our tool can be a worthwhile supplement to video portals such as *Qik* or *YouTube*.

Figure 2. Dummy prototype for customizing content demand

5. CONCLUSIONS

Our concept is intended to address the deficiency characteristic of ordinary video portals that provide large quantities of content but lack sufficient quantities of matched metadata. We want to hand over a tool to the users that encourages them to tag and rate audiovisual content by means of quick and easy questioning, in order to obtain requested entertainment media.

The users should be relieved of the burden of constructing excessively specific search terms to retrieve suitable content. The effort necessary to gain rapid access to an interesting video clip category should lie between that characteristic of established video portals (high effort, high level of customization: complex search by entering free terms) and that of a regular television program (low effort, low level of customization: instant access but no search capability).

6. REFERENCES

[1] Ames, M. and Naaman, M. 2007. Why We Tag: Motivations for Annotation in Mobile and Online Media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA, April 28 – May 03, 2007). CHI'07. DOI= http://doi.acm.org/10.1145/1240624.1240772

[2] Baluja, S., Seth, R., Sivakumar, D., Jing Y., Yagnik, J., Kumar, S., and Ravichandran, D., Aly, M. 2008. Video suggestion and discovery for YouTube: Taking random walks through the view graph. In *Proceedings of the 17th international conference on World Wide Web* (Beijing, China, April 21 – 25, 2008). WWW'08. DOI= http://doi.acm.org/10.1145/1367497.1367618

[3] Bellekens, P., Houben, G.-J., Aroyo, L., Schaap, K., and Kaptein, A. 2009. User model elicitation and enrichment for context-sensitive personalization in a multiplatform TV environment. In *Proceedings of the 7th European Conference on European Interactive Television Conference* (Leuven, Belgium, June 3 – 5, 2009). EuroITV'09. DOI= http://doi.acm.org/10.1145/1542084.1542106

[4] Chi, E. H. and Mytkowicz, T. 2007. Understanding Navigability of Social Tagging Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA, April 28 – May 03, 2007). CHI'07.

[5] Guy, M. and Tonkin, E. 2006. Folksonomies: Tidying up Tags? *D-Lib Magazine*, 12(1), January 2006. Available at: http://dx.doi.org/10.1045/january2006-guy (accessed 15 April 2010).

[6] Lerman, K. and Jones, L. 2007. Social Browsing on Flickr. In *Proceedings of the International Conference on Weblogs and Social Media* (Boulder, Colorado, USA, March 26-28, 2007). ICWSM'07.

[7] Mei, T., Yang, B., Hua, X.-S., Yang, L., Yang, S.-Q., and Li, S. 2007. VideoReach: an online video recommendation system. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (Amsterdam, The Netherlands, July, 23 – 27, 2007). SIGIR'07. DOI= http://doi.acm.org/10.1145/1277741.1277899

[8] Sen, S., Lam, S. K., Rashid, A., Cosley, D., Frankowski, D., Osterhouse, J., Harper, F. M., and Riedl, J. 2006. Tagging, communities, vocabulary, evolution. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work* (Banff, Alberta, Canada, November 04 - 08, 2006). CSCW'06. ACM, New York, NY, 181-190. DOI= http://doi.acm.org/10.1145/1180875.1180904

TV as an interactive medium to reflecting Energy consumption in daily life

Tobias Schwartz | Johanna Meurer | Matthias Beztz
FIT – Fraunhofer Institute for Applied Information Technology
User-Oriented Software Engineering
Schloss Birlinghoven
53754 Sankt Augustin Germany
tobias.schwartz | johanna.meurer | matthias.betz @fit.fraunhofer.de

ABSTRACT
In the last few years, a growing awareness has emerged on the impact that personal energy consumption have in the global levels of energy waste. Building on this trend, SmartMetering has focused on providing access to energy consumption data with increasing levels of granularity. In this position paper we show the results of a preliminary study of SmartMetering technology in different contexts of use. One repeating pattern observed in our studies was the relevance of the connection between metering data and usage habits. We argue that the efforts of SmartMetering are incomplete if they stop on data visualization without further supporting the change of habits in energy consumption. To illustrate our ideas, we provide a set of design sketches that move the focus from energy consumption displays towards supporting changes in personal and social behavior patterns which are ultimately the source of energy consumption.

Keywords
TV, Energy awareness, Smart Metering, Social Impact

INTRODUCTION

A principal goal is to minimize the deployment of resources to improve environmental protection by supporting sustainable products or process to prevent climate change. Many activities in industry, domestic context and research have been initiated in the last years with the key aspect to pollution control.

Monitoring and controlling energy usage is an important part to reach that aim and concerns consumers and the used infrastructure in equal measure. One feasible solution is to aim on the utilization of energy-saving appliances and infrastructures as a rule. This is often accompanied by high planning efforts and an appraisal of profitability.

Another feasible solution is to provide adequate information of energy usage to the consumer. Individuals could support to reduce their carbon footprint significant by changing their behaviors at home. Most of the energy used in a domestic context is invisible to the consumers. For unveiling unnecessary energy consumption it is necessary to provide feedback.

In this paper we propose a first theoretical approach to provide adequate feedback for monitoring and reflecting Energy Consumption "in the loop" by using TV as an interactive and omnipresent medium.

APPROACH

In this contribution we present the results of a preliminary study that supports this reflection, together with two complementary approaches for bringing personal behavior and energy consumption together to enable consumers to reach higher levels of energy efficiency. We illustrate these two approaches by means of a simple design sketch.

Smart Metering in Domestic Context

We conducted an empirical study in private households. We equipped 13 voluntary households with simple smart metering technology to measure energy consumption at device level in four different categories. Prior to the installation of the metering technology, we conducted semi structured interviews focused on the individual energy consumption of each participant. After one week of measuring, a collaborative data interpretation session was conducted with each of the participants.

Common ethnographic methods like site visits, participatory observation, interviews and workshops were used to enrich the studies. The gained empirical material was analyzed with qualitative methods of social science [1,2,3].

Supporting Change of Habits

In design sessions hold after the studies, we discussed approaches for supporting the change of habits. The results emerging from our studies clearly showed that the interplay between energy consumption data and personal habits was the key for stimulating energy efficient behavior. We present here two contrasting design concepts that illustrate a change of focus in the construction of energy efficiency fostering technology, from data visualization to habit change support.

TV as an omnipresent medium in domestic context

During the roll out of the smart metering sensors in private households we also ask the potential users of adequate feedback channels. In the most cases the TV as an omnipresent medium in private households were entitled. Regarding to that, we started to implement design sketches to illustrate energy consumption on TV. Thereby we

focused on a good energy transparency in situ and derived designs to develop different spaces of values and interaction mechanisms to support energy management on TV as shown in figure one.

Figure 1:Illustration of device clusters on TV

The development approach 'Living Lab' as an user centered design process aimed on a usable energy solution. In future work we will evaluate and improve the visualization on energy consumption on TV in field studies and participatory design workshops.

Figure 2: Energy consumption by priority as an current value

Improving the capture of behavior

One option to include behavior is to capture and tracking the personal activities and integrates this information with the energy consumption data. Like stated in the introduction, personal behavior is very complex to model but weak structured approaches could support users reconstructing their behavior in the past for an ex post reflection and analysis of their energy consumption. One possible implementation could be the introduction of a tool like the sensecam [4] to capture and document situations. This photo streams could be then synchronized with SmartMetering information and may help users to recall certain situations and reflect their in-situ behavior. This approach would allow the construction of histories which form the basis for ex post analysis to stimulate learning and reflection and motivate change of behaviors in the future [5].

Provide energy consumption information in situ

Reconstructing context could prove very difficult. A supplemental approach could be to provide information of energy use in-situ, during use. The situation is then enriched by direct feedback of current consumption. Energy use produces a breakdown in the activities of the users which motivates a reflection and has the potential of triggering a learning process. A possible implementation of this approach is the use of haptic or acoustic feedback responding to current consumption, or to changes in patterns of consumption.

Future work

In the future we intend to build prototypes of the concepts presented and evaluate them in both contexts presented here. Then we want to compare both approaches regarding their effects on a more energy efficient behavior. The results will provide interesting insights in the field of IT support for energy-efficiency. Beyond that, the results will be discussed against existing paradigms in the field of context-awareness, behavior sciences and e-learning

REFERENCES

[1] Tashakkori A, Teddlie C. Handbook of Mixed Methods Social and Behavioral Research. Sage Publications, Inc; 2002:784. Available at: http://www.amazon.com/Handbook-Methods-Social-Behavioral-Research/dp/0761920730.

[2] Kelle U. Sociological Explanations between Micro and Macro and the Integration of Qualitative and Quantitative Methods. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research. 2001; 2(1). Available at: http://www.qualitative-research.net/index.php/fqs/article/view/966.

[3] Wixon D. Qualitative research methods in design and development. Interactions. 1995; 2:19-26.

[4] Sellen AJ, Fogg A, Aitken M, et al. Do life-logging technologies support memory for the past?: an experimental study using sensecam. Conference on Human Factors in Computing Systems. 2007. Available at: http://portal.acm.org/citation.cfm?id=1240624.1240636.

[5] Chalmers M. A Historical View of Context. Computer Supported Cooperative Work. 2004; 13(3). Available at: http://portal.acm.org/citation.cfm?id=1039829#.

Enable Users to choose from different Channels – Using Multichannel Technology to enrich the TV Experience

Sebastian Groß
Pixelpark AG,
50676 Cologne
sebastian.gross@pixelpark.com

Sabrina Josellis
Pixelpark AG,
50676 Cologne
sabrina.josellis@pixelpark.com

Roland Westermaier
Pixelpark AG,
50676 Cologne
roland.westermaier@pixelpark.com

ABSTRACT
Within this paper we describe an IPTV prototype created by our company based on Microsoft Silverlight technology. It is intended to give a perspective on the range of opportunities brought about by interactive TV over IP and producer-consumer interaction using new technologies. Offering several additional channels from which the user can chose will enrich the TV experience. We also implemented two different modes of using the prototype. Both, a lean-forward (PC-version) and a lean-back (TV-version) scenario were realised.

Categories and Subject Descriptors
H.5.1 [**Multimedia Information Systems**]: Methodology

1. INTRODUCTION
Our company has developed a staging concept and consumer system in the form of an IPTV prototype based on Microsoft Silverlight. The professional user needs a staging concept as a basis for interactive real time storytelling about the live media event, in this case for the Olympic Games 2008 [1, 2]. However, this requires the creation of an interactive platform for the consumer, i.e. the viewer.

The staging concept provides a further base for communication and interaction between the professional user (producer) and the consumer. This determines the content that can be supplied to the consumer and the options for interaction. As such, the staging concept represents the foundation for interactive storytelling. Our company created an initial prototype based on the Windows Presentation Foundation (WPF) in order to underline this. This is followed by a second Web-TV prototype based on Microsoft's Silverlight technology. Results from a field trial were considered in a new version of the Olympic Games Live Staging Demonstrator, which briefly is described in this document.

In the following we present a brief overview of the format, the structure and the used contents within the prototype. Furthermore, the features and functions of the application will be explained. Last but not least we provide a list of possible additional features and an outlook to the future.

2. CONCEPT
The main feature of our IPTV prototype is the opportunity to enable users to choose from a multiplicity of several channels. For the prototype a number of main sport channels were chosable. In addition, video material could be added to complement the main channel. Both, a lean-forward (PC-version) and a lean-back (TV-version) scenario were realised. On the one hand this allows viewers to use their TV remote to navigate quickly and easily in full screen mode. On the other hand it supports the PC user who "leans forward" and uses a keyboard and mouse to navigate round the application in his browser. The already used content for the Main Sport Events including the sub channels are based on the material broadcasted during a field trial. More images and texts were added to this video content.

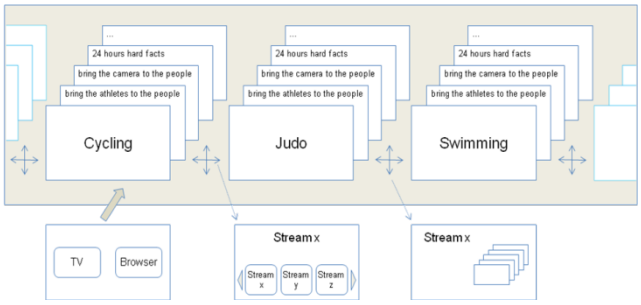


Figure 1: Multichannel concept

3. INTERFACE
The prototype focuses on moving images and the TV experience. As a result it is necessary to use a simple and intuitive form of navigation that can be used with a remote control or a mouse/keyboard.

Figure 2: Interface (user can chose between lean-forward and lean-back scenario)

The viewers can use the right and left arrow keys or their mouse to display all available **main sport channels** as an overlay of thumbnails in the lower area of the screen. The active channel is also displayed as a thumbnail. When pressing the OK button, the viewers get to the corresponding channel or, if they do not select anything, the overlay disappears again.

If the viewers want to access the **additional sub channels** they can use the up and down arrow keys or their mouse to display all available sub channels as an overlay of thumbnails in the left area of the screen. When pressing the OK button, they get to the chosen channel.

Additional information (such as videos, image galleries, texts with or without images, votes, ratings, etc.) is available for each video and is embedded using menu cards. The viewer presses OK to activate the card view. They can then use the arrow keys or the mouse to browse and select the additional information.

The **menu** consists of cards containing certain information that are piled on top of each other and an information area that is directly linked to the cards. A template that reads data from XML and creates content dynamically was developed for each type of information. When viewers browse the cards, the element whose content they select is highlighted in the list view. The list view of the information area is designed to provide an immediate survey of the available information. It also makes orientation easier. As before, the viewers use the arrow keys on the remote control or a mouse to navigate. They select a menu item by pressing the OK button or by clicking on a card.

Not just the additional content can be placed on top of the LIVE screen. The director can continue to influence events by using the **picture-in-picture function** to tell a story or highlight links to other channels. Therefore, the picture-in-picture function is the most important tool a director has for successful storytelling.

4. FURTHER USE

Further developments related to the SocialMedia topic are feasible. In the following chapter we describe some of the most interesting functionalities that may be considered for further work on that domain:

Chat – This would allow viewers to chat about the event while it is being shown on the screen.

Forum – A forum allows consumers to provide their own individual feedback. Qualified feedback in the form of contributions could be collected and evaluated as an extension of the rating.

Community – Community features that allow collaboration and the exchange of information between interested parties and pressure groups.

Rating – It could be possible to rate moving image content. Viewers can award a stream between one and five stars in a rating. The data is collected on a server and evaluated. The average rating is shown for each video.

Video-voting – Another variant of voting is the voting for videos. This allows the consumers to interactively influence the camera view so they can see what they want to see. Viewers can select from a number of videos and cameras.

5. ACKNOWLEDGMENTS

The Silverlight prototype is another step towards achieving a TV experience (at least as far as the TV scenario is concerned). However, the inclusion of the browser also means that it has the potential to reach a wider audience and can be viewed as a vision of the future.

6. REFERENCES

[1] Löffler, J., Güntner, G., Horstmann, H., Pereira, J. (2009): LIVE Staging the Olympics - A Novel interactive TV format and its supporting technology. Salzburg Research, ISBN 978-3-902448-18-7. August 2009.

[2] www.ist-live.org

Understanding Mobile Social Networks

Benjamin Mangold
Institute for Media Research
University of Siegen
57067 Siegen
+49 (0) 271 / 740 2289
mangold@ifm.uni-siegen.de

ABSTRACT
The paper presents and discusses a prototype for the spatio-temporal tracking and visualization of mobile user experiences and interactions with mobile media in mobile social networks. The tracking prototype consists of GPS-devices and mini video cameras. The approach is based on time-geographic methods of visualization. It represents user interactions, mobile communication and media use in its locative and temporal dimensions. The spatiotemporal approach to mobile communities allows insights about the structure of mobile interactions, mobile communities and mobile user contexts. It helps understanding social use of mobile media devices in social networks.

Categories and Subject Descriptors
H.5.2 [**User Interfaces**]: User-centered design

General Terms
Design, Documentation, Human Factors

Keywords
Tracking, Visualization, Mobile User Experience, Mobile Social Networks, Mobile Communities, Time Geography, Spatio-temporal dynamics, Human Centered Design

1. INTRODUCTION
The mobile phone has been a social device from its very beginning. It serves the maintaining of social ties and the management of social activities via mobile communication. With the integration of computation in mobile devices and the ongoing proliferation of mobile media applications, the mobile phone becomes increasingly relevant for the design of social media. For the design and planning of mobile media, understanding the situation and context of media use is a crucial factor. Interactions with mobile devices are highly dependant on the situational context and the actual user practices vary according to the socio-cultural situation of use. However, especially in the design of mobile media it is difficult to make clear judgments a priori about the situation of use, as the user is mobile and therefore the context is highly variable in mobile settings [1]. Recent research has tried to make the situation of mobile user experience plausible looking particularly at the locative dimensions of mobile media [2]. This approach is helpful for the design of locative media or location based services (LBS). However, this paper argues, that to fully understand the role of the situation, the temporal dimension of mobile use has to be integrated into research. Temporal regularities and routines in everyday life show to have a high effect on media use, as showed in Zerubavel's study of sociotemporal orders [3]. This shows to be especially relevant when looking at group and community interactions with mobile devices. To look at mobile user experiences in a spatio-temporal perspective could help to get a clearer understanding of social dynamics and the role of communication and media use in group interactions. This approach could be applied in research on mobile communities and mobile social network services [4]

Information about the situational context of use and group interactions are also highly relevant in the fields of design of emergency communication as a part of crisis response and crisis management. Emergencies are highly fluid events in which the relief forces have to continuously make sense of an unfolding event [5]. Especially in the early phases of emergencies or disasters, the knowledge about the situation is very limited. Mobile media and IT can support the establishment of a common sense about a situation, which is the basis for accurate assessment and adequate action.

This paper presents experiences with a prototype for tracking mobile media use in mobile communities. The goal is the development of a model for the spatio-temporal visualization of mobile media usage and mobile communication in a Geographical Information System (GIS). The model integrates locative and temporal dimensions of media use based on a time-geographic framework [6]. By understanding the spatio-temporal structures of mobile media use, particularly where (i.e. in what location) as well as when (i.e. at what time) mobile media are used, usage situations and structural patterns of mobile user experiences can be evaluated.

The highly dynamic work of emergency relief forces offers a good testing background for the evaluation of the model. The spatio-temporal model of visualization was therefore evaluated by analyzing the interactions in a fire brigade by tracking an emergency exercise. This research is part of the research project *Security Communication* at the University of Siegen, Germany [7].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
EuroITV□□□□, June 9th-11th 2010, Tampere, FINLAND.
Copyright 2010 ACM 978-1-60558-831-5 /10/06…$10.00.

2. METHOD

2.1 Time Geography

Background for the model is Time-Geography [6]. The basic idea is to display the movement and interactions of individuals on a 2D coordinate plane (with the x- and y-axis forming the longitude and latitude coordinates, neglecting the geographical height of the landscape) and with the z-axis representing the time axis. In this model a person standing still at the same place would generate a straight line growing straight up. This spatio-temporal "aquarium" has been used for the evaluation of social dynamics as well as the analysis of the distribution of activities in space and time [8, 9, 10]. Here this theoretical framework is used to represent interactions with mobile media (e.g. radio) and interpersonal communication in a space-time cube. It therefore offers the background for describing an unfolding situation and the positioning of communication and media usage in space and time.

2.2 Prototype

The prototype for the spatiotemporal tracking currently consists of ten tracking sets made up of a GPS-device and a mini-camera that can be easily attached to a person's clothes. The video equipment currently allows up to 1h 45min of continuous recording of audio and video in VGA quality (640x480 pixels). The recording limit results from camera software limitations. The overall equipment is of very small size, so it is not obstructive to the person wearing it.

3. FIELD EXPERIMENT

The prototype was tested in a one-day emergency exercise of a local fire brigade in Rietzel, Germany. For this field experiment eight firefighters were equipped with the GPS/Camera prototype. There were about twelve firefighters in the field, acting in up to three different units. Therefore the eight tracking prototypes gave a nearly complete overview of the unfolding situation.

Figure 1. Two video screen captures of firefighter camera

The exercise consisted of a coordinated effort at different places in a village and its near surroundings. First was securing and controlling a large wood fire (which for training purposes had been ignited by the firefighters themselves), second the removal and chopping of a tree in the village center. Furthermore a variety of tasks had to be done on short notice distributed at several places around the village, as e.g. securing a water source and getting new materials. By distributing the participants of the exercise, communication and coordination of groups via radio could be trained. For the execution of the tasks the fire brigade had to split up in different smaller units, which gave the opportunity to not only analyze interpersonal voice communications, but also take a look at radio communications and its relation to the actions taken.

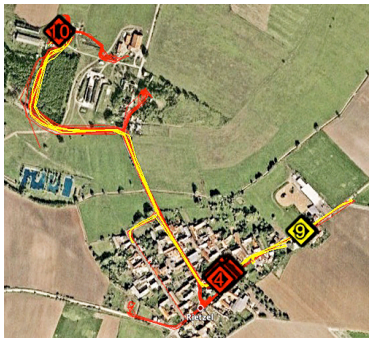


Figure 2. GIS visualization of the fire brigade exercise

3.1 Data gathered

The experiment showed that the quality of the material gives a good impression of what the person is currently doing and seeing. It furthermore allows an understanding of the wearer's voice, the voices of people in his surrounding and the radio communications. The audio data was still understandable when the firefighter was using breathing equipment. The material could therefore be easily transcribed and analyzed.

The GPS tracks provided a good accuracy of about 5m in the open field. In buildings, the connection to the satellites was lost or not very accurate (variations up to 20m). Still, being provided with the entry and exit places of buildings and the time of entry, in connection with the video stream of the firefighter, the activities and movements in buildings could be reconstructed.

3.2 Results

The acceptance of wearing the camera by the firefighters was high, in fact it seemed that the wearers forgot about after fifteen minutes into the exercise that they were wearing the tracking equipment. After the exercise the eight tracks of video and GPS made it possible to assess and evaluate mobile user contexts and mobile interactions from multiple perspectives. Especially problems in radio communication or misunderstandings in coordination efforts between distant groups could be analyzed for all communication participants individually. Therefore the prototype showed to be a promising basis for further empirical research on the design of mobile media and human interactions with mobile devices in social contexts.

4. DISCUSSION AND OUTLOOK

The field experiment was an opportunity to test the prototype for the tracking and the visualization of mobile user experience. The results so far seem to be a promising basis for a better understanding of mobile user contexts, media interactions in mobile contexts, the positioning of interactions in space and time and the dynamics of interactions in mobile communities. Still the prototype has to be iterated and tested in other community settings.

The prototype could be applied in the field of crisis response and management, e.g. for the tracking of emergency operations and the evaluation of training exercises.

The observations also serve the design of mobile media technologies, as the user interactions can be reviewed and evaluated from multiple user perspectives. The possibilities of the

tracking equipment therefore also promise fruitful applications in mobile and social media design and Human Centered Design.

Upcoming research challenges will be questions of scope of the tracking (what would be the right size for this kind of research), data management and the optimization of the evaluation workflows (e.g. how interviews and personal observations could accompany the tracking), as well as issues of privacy.

5. ACKNOWLEDGMENTS

Sincere thanks go to the fire brigade of Rietzel, Sachsen-Anhalt, Germany.

6. REFERENCES

[1] Cherubini, M. and Oliver, N. (2009) A Refined Experience Sampling Method to Capture Mobile User Experience. CHI 2009. Boston

[2] Döring, J. and Thielmann, T. (eds.) 2008 Spatial Turn. Das Raumparadigma in den Kultur- und Sozialwissenschaften. Bielefeld: transcript.

[3] Zerubavel, E. (1981) Hidden Rhythms. Schedules and Calendars in Social Life. Berkeley, Los Angeles, London. University of California Press

[4] Lugano, G. 2008 Mobile Social Networking in Theory and Practice, First Monday, Vol. 13, No. 11.

[5] Landgren, Jonas 2007 Designing information technology for emergency response. Doctorial dissertation. Göteborg University.

[6] Hägerstrand, T. 1970 What about people in the regional science? In: Papers in Regional Science. Vol. 24, No. 1

[7] Sicherheitskommunikation. Institute for Media Research, University of Siegen. http://www.sikomm.org/

[8] Shaw, S. 2009 Towards a GIS-based Analytical Timer-geographic Framework for Physical and Virtual Activities. http://web.utk.edu/~sshaw/NSF-Project-Website/

[9] Ren, F. and Kwan, M. 2007. Geovisualization of human hybrid activity-travel patterns. Transactions in GIS, Vol. 11, No. 5, 721-744

[10] Calabrese, F., Reades, J., and Ratti, C. 2010 Eigenplaces: analysing cities using the space-time structure of the mobile phone network. IEEE Pervasive Computing, Vol 9, No. 1, 78-84

Ambient-Aware Service Infrastructure for Home IT Environments

Benedikt Ley
University of Siegen
Information Systems and New Media
Hoelderlinstr. 3, 57068 Siegen
+49 271 740 3265
benedikt.ley@uni-siegen.de

Martin Stein
University of Siegen
Information Systems and New Media
Hoelderlinstr. 3, 57068 Siegen
+49 271 740 3382
martin.stein@uni-siegen.de

ABSTRACT
In this paper we describe an approach for an ambient-aware service infrastructure for home IT environments. To make high functional home electronic devices accessible and manageable for users, we developed a framework to make local available services in domestic environments visible und usable on the user's mobile devices. To enable this dynamic process we used a high flexible integrated approach of component-based and service-oriented architectures.

Categories and Subject Descriptors
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Design, Human Factors

Keywords
Home IT, ambient-awareness, service infrastructure, OSGi

INTRODUCTION
The ongoing digitalization of the media environment affords a variety of changes to home IT. Today many TVs, e.g. in combination with set-top boxes, already include network connections with access to the Internet. This trend is also recognizable on other home electronic devices and brings along opportunities for new services and features. As a consequence, the changing of use behavior requires new user interfaces to make these services manageable for users. In this work we present a framework that provides a dynamic service infrastructure for domestic IT environments. For this, home electronic devices make their services accessible to nearby users who can install and use them on their mobile devices. Since an increasing number of services will cause highly complex and overloaded interfaces, we additionally integrated an ambient-aware approach to remove or hide services that are irrelevant within the nearby environment.

STATE OF THE ART
In order to react quickly to changing user requirements, software development processes need to be efficient and flexible. That is why paradigms like component-based or service-oriented architectures (SOA) became so very relevant within the last few years. The abstraction of services and components enables not only software developers but also end users to extend or adopt existing software systems in an efficient and dynamic way. There are several approaches that can be used in order to utilize the characteristics of respective paradigms in relation to specific issues. [2] shows an approach of combining component-based systems with SOAs by adapting a component-based system according to available services. In the area of mobile computing [6] shows how to find and interact dynamically with services during runtime.

A combination of multimedia applications with component-based systems is shown in [5]. The authors integrate MHP (Media Home Platform) into a component based OSGi (Open Service Gateway Initiative) environment to allow the control of several home electronic devices within MHP.

Performing ambient-based software adaptations, software systems need to be aware of the current environment. For this, several approaches and existing systems are using the GPS technology that certainly does not work within indoor surroundings. The existing works in [3], [4] and [1] address this issue by gathering local Wifi, ZigBee or RFID information and derive current location information from it.

OUR APPROACH
In this work we want to describe a rich and integrated solution for the usage of Home IT. Because of the heterogeneous infrastructure of current devices, it is important to create a platform-independent framework to enable the integration of several electronic devices. To develop such a dynamic framework we need a flexible architecture to adopt and extend the system regarding the changes of ambience during runtime. For this, we use an integrated approach of component based architectures and SOA by applying the role model of SOA, consisting of service broker, service provider and service consumer, but using locally available components instead of web services. Through this combination we are able to achieve a cross-platform communication and functionality without static dependencies.

Because a change of the location means in most cases a total change of the technical infrastructure (e.g. living room vs. kitchen), the functionality offered by a system to meet the requirements of interacting with any available device would cause highly complex and overloaded interfaces. This leads to bad user-experiences or even exceeds the capabilities of handheld devices in the manner of input space or interaction effort. To overcome this issue our approach tries to utilize ambient-information about a location and its devices to offer just ambient-relevant services to users. The services will be made available when entering a location and will be provided through components and they will be hidden or removed from the device when leaving the location.

The local installation of components enables the device to provide some functionality without being dependent on any other device as well as an efficient way to communicate with other devices because of a smaller message overhead. Besides the ambient-aware services there are also services that can be used permanently and will remain on the device independent of the user's location.

For some reasons it might be important to ensure that users will download and install some components. Therefore we distinguish between three kinds of service components:

- Mandatory: component must be installed (e.g. security update or research evaluation survey)
- Recommended: component is recommended in the current ambience (e.g. remote control for a TV in a living room)
- Optional: component is available for download but unrelated to the current-ambience (e.g. video player on mobile device)

Based on these assumptions we developed an early prototype that is described in the next section.

IMPLEMENTATION

The prototype we developed consists of a Media Center PC, a server and a mobile device. The Media Center PC as a home electronic device can be used the server as well. Because of the limited bandwidth of Near-field networks we decided to implement a lightweight communication model for the server-mobile interaction using TCP/IP-sockets. The communication between server and Media Center PC is realized through web services to achieve a high level of independence. For the mobile devices we used an Android smartphone in combination with an OSGi framework. Thus we are able to abstract from the underlying configuration (OS, programming language, etc.). This enables us to port the developed applications to every platform that offers or is compatible with the OSGi standards. In addition the OSGi framework provides dynamic mechanisms for hot deployment, which means that new components and functionalities can be accessed without restarting the mobile device or the framework.

As a part of the prototype we developed bundles, which are responsible for the adaptation to different media infrastructures. Therefore they need to introduce the mobile device (service consumer) to other devices in the location. After broadcasting an entry-message to the environment, the mobile device will receive an XML-based description of all available components for the current environment from a server (service broker). The description contains meta-data (e.g. name, version) and the URL to the location (service provider) where the components can be downloaded.

Because the entry-broadcast will be published in the whole underlying network, the granulation of locations and therefore the amount of devices that are part of it, is limited by the used network technology. Currently we rely on Wifi as a communication medium, but for further development it is possible to use other techniques supporting TCP/IP and UDP/IP protocols like ZigBee or Bluetooth for more detailed ambient-awareness.

CONCLUSION AND FURTHER WORK

We presented an ambient-aware service infrastructure to make highly functional home electronic devices accessible and manageable for the users. These services can be installed and used on user's mobile devices during runtime. To perform this dynamic process we used an integrated approach of component-based and service-oriented architectures. We also involve ambient-information, so that only locally available services are visible and usable to the users. In our further work we want to evaluate this approach in about 20 selected households that are currently equipped with Media Center PCs and mobile devices. Furthermore for a useful evaluation we need to develop several example services.

ACKNOWLEDGMENTS

This work is funded by the Ministry of Economic Affairs and Energy of the State of North Rhine-Westphalia within the SocialMedia project.

REFERENCES

[1] Betz, M., Ley, B., Pipek, V., Wulf V. Folksonomies for real things: Tagging objects with RFID as a source for context-awareness. *Informatik □□□□ □Beherrschbare Systeme dank Informatik* (2008), 787-794.

[2] Cervantes, H., Hall, R.S. Autonomous Adaptation to Dynamic Availability Using a Service-Oriented Component Model. *Proceedings of the □□th International Conference on Software Engineering* (2004), 614-623.

[3] Cheng, Y. Using ZigBee and Room-Based Location Technology to Constructing an Indoor Location-Based Service Platform. *Intelligent Information □iding and Multimedia Signal Processing□International Conference on* (2009), 803-806.

[4] Luimula, M., Kuutti, K. Locawe: a novel platform for location-aware multimedia services. *Proceedings of the □th International Conference on Mobile and □bi□uitous Multimedia* (2008), 122-129.

[5] Díaz Redondo, R.P., Fernández Vilas, A., Ramos Cabrer, M., Pazos Arias, J.J. Exploiting OSGi capabilities from MHP applications. *□ournal of Virtual Reality and Broadcasting*. 4, (2007), 16.

[6] Sanchez-Nielsen, E., Martin-Ruiz, S., Rodriguez-Pedrianes, J. An open and dynamical service oriented architecture for supporting mobile services. *Proceedings of the □th international conference on Web engineering* (2006), 121-128.

DTG Digital TV Group

amea ambient * media * association www.ambientmediaassociation.org

iDtv

digibusiness.eu Gateway to Digital Finland

Deutscher IPTV Verband

informitv

The City of Tampere

EuroITV2010

9th June-11th June 2010, Tampere, Finland

www.euroitv2010.org

EMMi Lab. - www.tut.fi/emmi
Dept. of Business Inform. Management
Tampere University of Technology

NOKIA CONNECTING PEOPLE

Microsoft Mediaroom

DEMOLY®

digitampere.fi

NoTube This is future TV

YLE

amea ambient * media * association www.ambientmediaassociation.org

SIGCHI

acmsigmm SPECIAL INTEREST GROUP MULTIMEDIA

sigweb

acm