Lugmayr, A., Lyytinen, J., Mailaparampil, A., Tico, F., Maijala, S., Tuominen, S., Pihlajamäki, T., Rautavirta, P., Oksanen, M., Spieker, J., Niiranen, S. & Kalli, S. 2002. The future interaction tv project developing diet - digital interaction environment for tv. Proceedings of the 5th Nordic Signal Processing Symposium, NORSIG 2002, October 4-7, 2002, on board Hurtigruten, Norway. 6 s.

THE FUTURE INTERACTION TV PROJECT
DEVELOPING DIET - DIGITAL INTERACTION ENVIRONMENT FOR TV

Artur Lugmayr, Jussi Lyytinen, Anurag Mailaparampil, Florina Tico, Satu Maijala, Sami Tuominen, Timo Pihlajamaeki, Perttu Rautavirta, Mikko Oksanen, Jens Spieker, Samuli Niiranen, and Seppo Kalli

Signal Processing Laboratory, Tampere University of Technology,
P.O. Box 553, 33101 Tampere, Finland, {firstname.surname@tut.fi}
Contact Author: {lartur@acm.org}

ABSTRACT

Digital interactive TV (digiTV) has been deployed in several countries. Within the *future, interaction TV project (fiTV)* we look beyond the next five years and try to find new innovations, that might enhance this platform by new service types. Within this paper we provide an overview of our key-findings during the first two project phases, such as MPEG-21 compliant universal multimedia access (UMA) on digiTV, MPEG-7 metadata streaming facilities in parallel to audio/video, *Multimedia Home Platform (MHP)* [7] compliant software implementation, development of a service architecture, and harmonization of metadata standards to a digital broadcast item (DBI) utilizing the digital broadcast methodology (DBMI).

1. INTRODUCTION

For the duration of three years the futureTV project [8, 12], researched the basic distribution network and local facilities for the first phase of digital, interactive TV. It provided a general service deployment framework, and gave the way for future research directions, that might need further investigation. Therefore by the mid-2001 a new project was launched for the duration of three years: the future, interaction TV (fiTV – http://www.futureinteraction.tv) project.

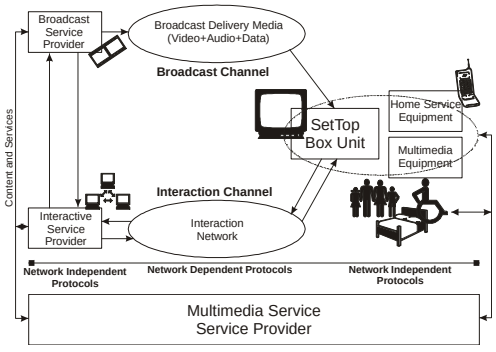

Figure 1. Generic service architecture for service provision on the digiTV platform.

The overall project is a cooperation between different partner universities (Tampere University of Technology – Digital Media Insitute, Tampere University - Hypermedia Laboratory, Helsinki University of Technology – Telecommunications Software and Multimedia Laboratory) in Finland, and major Finnish telecommunication companies. Within the scope of this paper, we concentrate on the services, as developed by the Digital Media Institute, Tampere University of Technology.

The overall project emphasizes several research interests within the service deployment chain, from service editing, -broadcasting, -utilization at consumer side, and involvement of different feedback architecture types. We rely that each set-top-box (STB) is connected to a home network, furthermore to a Internet connection, facilitated as feedback channel to multimedia service providers. Thus, the three key elements and their correlation essential to interactive digital television are on main focus: the underlying technology, consumer needs and implications, and different content types. We aim at standardization processes, finding new innovations, research gaps, and interaction models and by proofing our concepts with our implementations. The consumer shall move to a more active state, where he can actually "interact with digiTV content" [2, 13].

Several broadcasting related questions are addressed: security and content authentication, technical local facilities, evaluation of interactivity strategies, service authoring, interacting and navigation through content, multimedia content repository for storing digiTV content, tools and modeling, involvement of wireless devices [9], and services deployed within a distributed service pool.

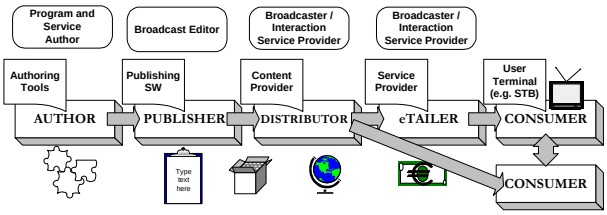

Figure 2. Principal digital broadcast chain.

Consumers of digital, interaction TV content do not want to pay more, be penetrated by more advertisements, and use more complex equipment. They expect simplicity and easy to use equipment. On the other hand, content providers goals are less production costs, relief of state governed regulations, and simpler production tools to win consumer's loyalty. To satisfy these needs our goals and questions are:

- *How can we bring services over the broadcast- or feedback channel over fully-automated service architectures?*
- *Which innovative interaction services in different domains (e.g. entertainment, education) are emerging in the next few years and beyond?*
- *How can we participate in standardization efforts, and develop strategies and pathways for the future based on our conceptual work?*
- *How can we satisfy our industrial partners by acting as a certain kind of test bed for their products?*

Three levels of research tasks are aspired to cope with those questions. General questions are addressed during the *Evaluation Level*, leading to proof-of-concept implementations covered by the *Development and Implementation Level*. The outcome of previous levels leads to the *Deployment and Integration Level*, focusing on the convergence of multiple end-platforms.

- *Evaluation Level - Technology & Content:* General questions and obtaining knowledge about current research and development work for finding innovative new research gap is aimed at.
- *Development and Implementation Level - Interactive Content & Concepts & Metadata:* Interacting with digital TV content and bringing the consumer on the TV screen are researched. The key is to satisfy the need of consumers to be seen on TV, to consume content, and to interact with it. The underlying technology is based on metadata, commonly referred to as data about data, and expressed by utilizing the extensible markup language (XML) is major entire part of this research phase. We aim at a harmonization of TV-Anytime [4], MPEG-7 [3], and MPEG-21 [1] standards, and a clear separation between metadata and referenced content. Our use-scenarios are outlined within the scope of this research paper.
- *Deployment and Integration Level - Future Interoperability:* Inter-device communication between wireless- (e.g. mobile phone), and wired (e.g. IP network devices) units are addressed by this project sub-task. This means the development of distributed home multimedia services and their convergence to a multimedia home entertainment center, accessible from digiTV equipment acting as service access point.

Within the scope of this research paper we review current achievements and results, closing by a short preview of the next project phase and its aspired goals. Several sub-tasks lead to the overall *Digital Interaction Environment for TV (DIET)*, an architecture upon which we deploy our visionary services in different application domains.

2. EVALUATING TECHNOLOGY & CONTENT

During the evaluation phase we considered questions and consumer-centered approaches of service development. We emphasized consumer-friendly services categorized according different topic clusters. From the technology point of view a topics such as development tools, secure access to services, current research works and business models, content personalization, and service architecture types for content distribution were addressed. In the following major research topics are stated.

2.1. Selling Content and Services

Even though our visionary and innovative services, they have to be sold in the business world. T-commerce is the new term for expressing digiTV content related business models. We summarized business models, secured and content authenticated payment schemes, and domains relevant for deploying digiTV services.

2.2. Review of International Research Initiatives

What happens in international research, and the evaluation of current research topics and their impact on digital television resulted in an identification of opened research issues, and continuous justification of scientific research directions.

2.3. Sensing Bits and Bytes (Digital Items)

Myriads of digital material on our devices without any intelligent mechanism for its organization lead the customer to loose any emotional and personal binding to them. The consumer tends to be content manager instead of content consumer. Awareness and personal relation to each piece of information requires approaches to perceive digital content in form of music, video files, multimedia presentations, etc. in an adequate and more physical, perceptive, and "real" way. Our evaluation leads to the approach of digital multimedia items (DMI), where each content entity (e.g. video) is annotated by metadata structures (e.g. content description of a video frame), defining its visual appearance. The basic starting point was the exploration of the capabilities of MPEG-21's digital item- (DI) [1], and digital rights management (DRM) declarations. We covered security and conditional access schemes, unique item identification, visual item representation, copyright protection, and UMA schemes to content from different consumer platforms (e.g. mobile, PDA, Set-Top-Box).

2.4. Perception of digiTV

How does the consumer perceive digiTV content and how it can be manipulated to obtain enhanced content "add-ons" is the major issue within this working topic. Our research work resulted in the formula "perception = virtual reality + immersion + interactivity + community"

as key to advanced content perceiving digiTV systems. Virtual reality methods create the virtual narrative space and determine the way, how content gets presented to the auditorium, by applying computer vision methods. How the consumer can handle and navigate within the presented content is covered by various interaction facilities. We classified interaction models according the degree of consumer involvement and available hardware.

- *Weak Interactivity* is based on a simple remote control;
- *Hybrid Interactivity* integrates input devices such as PDAs, web-cameras, speech interfaces;
- *Collaborative Interactivity* enables community creation (social aspect) and emphasizing social aspects;
- *Strong Interactivity* relates several communication channels, either between devices or consumers to each other;

Furthermore we developed an active content strategy, where video/audio/services are enhanced by metadata streams to manipulate content. One example is customization of actors, hyperlinked TV, exchange of premarked rectangular objects in a scene, and convergence with wireless protocol types, as referenced in [10, 2, 9].

2.5. The STB is alive: The Personal Comrade

The optimal delivery of the right content to the right consumers at the right time is a very future oriented task. Furthermore content shall not be searched by the consumer; it shall come to him according his needs and requirements. Intelligent digiTV equipment that observes itself, delivered content, and its surroundings enables giving suggestions and observing the consumer's mood and emotions to provide value added content in an appropriate way. For the lonely one's the TV shall be a "partner" or artificial "friend". The aim was the research of principles, issues, and evolutionary methods for realizing a digiTV box that talks, personalizes (e.g. based on intelligent agents), repairs, etc. A realization during the project might be too complex, but even though concepts and basic research work into this direction will be emphasized. As simple example of a use scenario: if the consumer falls "asleep" during a broadcast show, the TV should just simple switch off.

2.6. Club "Virtual Couch Potatoes"

Currently there is Human-Computer interaction, rather than Human-Human interaction. digiTV is a well accepted platform in the society, but there is still too much of a "Computer" interfaces. We evaluated digiTV for achieving more Human-Human interaction, and letting the "Computer" platform disappear as much as possible. Beside enhancing user interfaces, we focused on collaborative issues. The consumer wants to see himself during the broadcast show, and communicate with other ones. Therefore we developed communication strategies, and computer games, enabling bringing people together under the platform digiTV.

2.7. Usability and EduTV

Usability is a key question to move people to a state, where they can use our developed services in an easy way. To complex user interfaces scare consumers to use it. The topic is extended to educational TV, where the question of how digiTV can be utilized as distance education platform is followed.

2.8. Semantic digiTV

TV is a highly narrative medium, and tells different stories, which the consumer can enjoy passively. Especially digiTV provides the technological facilities for telling stories with different endings and parallel scenes.

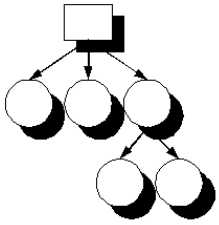

Figure 3. A parallel story with a definite starting point and multiple story flows with more endings.

Within the project we research how it is possible to fit different video/audio/services in a more human adequate form together for enhancing digiTV's narrative facilities. We try to build a ontology for an intelligent, narrative multimedia presentation to move the consumer form the passive "couch potato" position (prefiguration) - where he must interpret the presented content to the configuration state - where he can influence the presented topics and its flow or story evolvement.

3. INTERACTIVE CONTENT & CONCEPTS & METADATA

The evaluation phase gave freedom to more visionary and innovative thoughts and was followed by a more implementation oriented viewpoint. Within this section we focus more on realization and implementation concepts. Our key element for deploying services either over the broadcast- or feedback channel is the service architecture.

3.1. Service Architecture

The underlying hardware- and software architecture, where services are realized and tested convolves the provision of an infrastructure for storage, distribution, delivery, and service verification. Each party present in the deployment chain, thus service authors, publisher, distributor, eTailor, and consumer shall be present in the final unified setup for service distribution and exchange between several involved business partners: enhanced broadcasting architectures and its facilities, and the manifold feedback architectures and its rich protocol types.

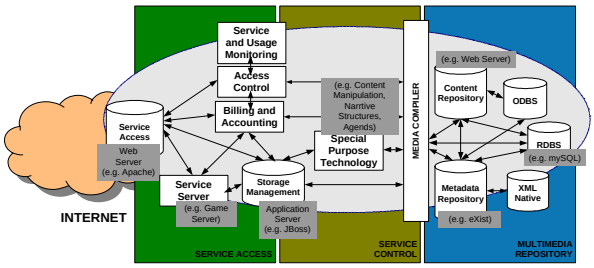

Figure 4. Feedback channel architecture.

During early project phases we realized a service architecture that utilizes the feedback-, and broadcast channel. Storage of digiTV related multimedia material, such as movies, applications, etc. requires the development of an intelligent multimedia repository for building an adequate distribution network for the next generation of digiTV services.

3.1.1. Feedback Architecture

A distributed service pool is the central element for providing access to value added content. Our approach toward the feedback architecture results in a three-tier model. The front-end (service access tier) convolved a commonly used web-deployment infrastructure, thus web-server, dynamic HTML pages, XML content servicing, etc. Certain business intelligence is enabled by the middle tier (service control), performing advanced tasks, such as content manipulation, business transaction management, back-end access control, etc.

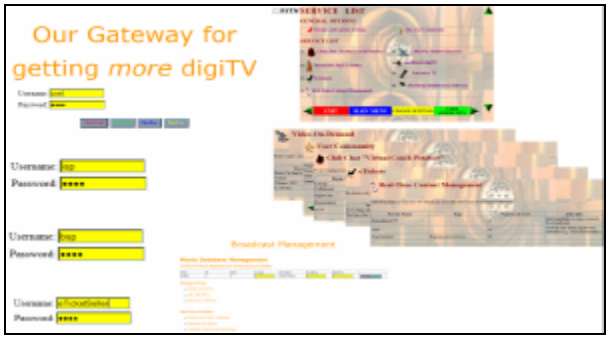

Figure 5. Login screens of our service architecture.

The major part of the back-end architecture is a multimedia repository, consisting of a content-, metadata-, and relational database for storing multimedia entities.

Besides the exchange of content between business partners our service architecture acts also as multimedia home server solution, for accessing multimedia content within the range of a household. The architecture is shown in Figure 4. Figure 5 shows the login screen of our feedback network.

3.1.2. Broadcast Architecture

We have a test environment comparable to real-world deployment architecture with current state-of-the -art digital broadcast equipment (audio/video multiplexer, modulator, application play out) including a DVB compliant STB.

3.2. Metadata based Services in DigiTV

Metadata based Services in digiTV are currently heavily researched, but it is still under standardization process and their usage and realization not defined. Within the project we tried to converge multiple metadata standards related to digiTV (MPEG-21, MPEG-7, TV-Anytime, SMIL) [1, 3, 4, 5, 6] and encapsulate different digiTV content types into one unified structure for presentation to the consumer. For exchange between several parties in the broadcast chain, and developed the framework for a generic, configurable DBI.

Figure 6. Hyperlinked TV for an ice fishing competition, where metadata description are used in the content overlying plane.

3.2.1. Electric Ticket Distribution

eTickets are the electronic form of tickets for leisure occasions. We heavily researched their deployment and mechanisms to deliver them from the ticket seller to the consumer. One major focus of research was security mechanisms (e.g. stamping patterns for eTickets) and their deployment on a digiTV network.

3.2.2. Hyperlinked (Segmented) Television

Hyperlinked TV is comparable to the WEB. The new “hyperlink generation” is used to surf through a net of hyperlinks. We try to merge both paradigms together for deploying hyperlink enabled services under the umbrella digiTV during a broadcast show.

Each video content type consists of movie segments. For each segment a representative frame is selected and annotated with MPEG-7 descriptors. Those descriptors contain information about the position of hyperlinks, action definitions for performing arbitrary task upon selecting one link, and image text components for textual hyperlink description. During a broadcast the consumer

gets notified of transmitted hyperlinks, and can navigate through them and select its content.

3.2.3. Digital Broadcast Item Methodology (DBIM)

Currently we work on this methodology, providing a unified framework for the delivery of digital TV content, and metadata. The consumer shall perceive the content as one entity to different terminals, comparable to a HTML page. In the near future we hope to present our current results of our methodology.

3.2.4. Agent Based Personalization

Personalization is required as people are overwhelmed with content. digiTV will provide a multiplication of new content types and services. Therefore a contemporary approach is required to find new techniques to filter the content that the consumer would like to experience. We researched one solution based on agent technology in close cooperation with the Health Care TV project (hctv.dmi.tut.fi).

3.3. DigiTV Client Development

3.3.1. HTML Browser

In close cooperation with the HCTV (Health Care TV Project hctv.dmi.tut.fi) we develop a HTML browser for a digiTV compliant consumer device. It shall obey similar, but more restricted features, as its PC based counterpart.

3.3.2. Client Architecture - eGames

eGames are connecting people and should provide a community for bringing people together on the digiTV platform. User interface independent development is our major concern to enable service delivery to multiple end-devices; therefore we develop and abstract modules for PC, STB, and PDAs. Our approach was entirely bottom up, for obtaining a software library of components reusable for future application development. Our current modules consist of a network- (udp, tcp, smtp, pop), xml parser-, graphic rendering-, virtual keyboard-, and user interface components.

3.4. Interaction Models

The consumer shall be independent from any type of input devices, such as a remote control. He shall navigate, and interact with content with a vision-, or audio based user interface. Our approach is based on a PC based home-server solution for performing calculation intensive tasks, where the digiTV equipment acts as client for handing transmitted input data (e.g. position of hand, emotion). The equipment itself can decide which actions to take. Our current implementation, a WebCam based user-interface, acts as virtual mouse. A marker enables the consumer to move a mouse cursor over content, and select regions. Currently we apply this technique to hyperlinked television as new input device for selecting links in the broadcast show.

Figure 7. Universal multimedia access (MPEG-21) based TicTacToe game for PC and STB.

Figure 8. Service Navigator.

3.5. Business Models

New Business models are required for new innovative technologies. Their research and investigation led to new findings and new earning types developed for our industrial partners.

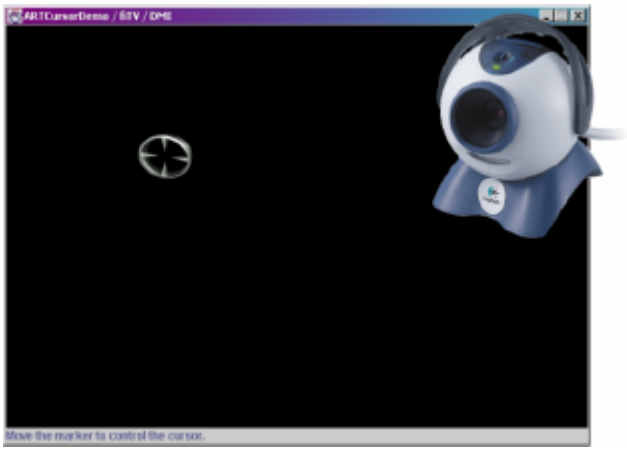

Figure 9. Interaction model: WebCam user interface (vision based digiTV mouse).

4. FUTURE INTEROPERABILITY

Throughout the project a continuous collection and revision of research perspectives is aimed at, due to rapid changes of the digiTV marked. Therefore, based on our two previous phases' results, future interoperability concentrates on integration aspects of digiTV equipment. Within this section we revise our goals during the next project phase.

4.1. Platform Independence

The entire goal is the realization of a solution for providing universal multimedia access to digiTV content anywhere, anytime, and anyhow. We will utilize PDAs, PCs, STBs, and wireless devices (3G) to proof our conceptual works. We are focusing on services, rather than audio/video content. The same profiled service shall be available across platforms, where high performance consuming task shall be performed at server side, rather than on less powerful consumer devices. We currently develop a platform independent chat application, to verify our concepts.

4.2. Metadata Convergence & Semantic

Still there are multiple metadata standards as based on XML in the world of multimedia. Basic research has been done already in previous research levels, but will continue throughout the project. As basic application we considered a video-preview and selection application, where content can be previewed based on metadata descriptions and sample strips. The consumer can select each and buy the complete structure upon preview. Another main issue is personalization, and how content can be personalized by including semantic information.

4.3. Intelligent Interaction with Content

The goal is to extend the current approach of intelligently interacting with content by emotion based user interfaces, utilizing vision-, and speech input. Emphasize is put on the relation between consumer, and 3D content. Emotion based user interfaces, utilizing vision- or voice. Especially for people with special needs it is an entire requirement to facilitate enhanced interface types. We hope to be able to present more results after finalization of the next project phase.

4.4. Security and Digital Rights Management

The current implementation of an eTicket exchange service concentrated on secure transmission, and sophisticating stamping pattern mechanisms. Besides, the security concerns can be extended to digital rights management of content. The consumer of digiTV content shall not receive it for free, and copyright protection rules shall be introduced. Our approach facilitates the MPEG-21 standard as starting point to digital rights management.

4.5. EdutainmentTV

Edutainment on digiTV, as domain related appliance of our current services covers mainly evaluation and conceptualization work. One approach is the integration of the learning object model (LOM*)* into the world of digiTV. Currently the LOM is only deployed on the web-platform. A case study shall lead to user-interface design screenshots, or sketches, serving as basis for a future system realization.

5. CONCLUSION & ACKOWLEDGEMENTS

But still, more research work is required. During the next phases of the project we will focus on service types, such as advanced narrative structures, edutainment TV, interconnection of digiTV equipment and home services, and enhanced interactive game structures on the platform digiTV in close cooperation with our project partners.

We would like to thank all our research colleagues at the Digital Media Institute, especially at the Signal Processing Laboratory, for their help and cooperation. Credits also go to our project partners at the Helsinki University of Technology, and the Hypermedia Laboratory at the Tampere University. Artur would like to thank Olli Stroem and Pasi Rustanius for teaching him a "bit of" ice-fishing.

REFERENCES

[1] ISO/IEC, *MPEG-21 Information Technology, Multimedia Framework*, Part 3, Digital Item Identification and Description.

[2] A. Lugmayr, A. Mailaparampil, P. Rautavirta, M. Oksanen, and S. Kalli. Applying MPEG-21 in Digital Television: Use Scenarios: ePostcard, eTicket, and eGames. accepted, IEEE Multimedia. ICME 2002.

[3] ISO/IEC JTC1/SC29/WG11, *MPEG-7 Standard and Reccommendations*, Pattaya, December, 2001.

[4] TV-Anytime Forum, *TV-Anytime Standards and Recommendations*, http://www.tv-anytime.org.

[5] ISO/IEC JTC1/SC29/WG11, *MPEG-21 Reccommendations*, http://ipsi.fhg.de/delite/Projects/MPEG7/Mpeg21.html

[6] WWWC (W3C), www.w3.org.

[7] Digital Video Broadcast (DVB), *MHP: Multimedia Home Platform Specification*, Vers. 1.0.2 Rev. 2, European Broadcasting Union (ETSI,EBU), Februar, 2002.

[8] A. Lugmayr, et. al, "The FutureTV Project: MHP Compliant Software Development," *FINSIG 2001*, Proceedings of, 2001.

[9] A. Lugmayr, et. al, "DVB over Wireless Networks," *Multi-Dimenional Mobile Communications*, Proceedings of, Pori 2001.

[10] A. Lugmayr, et. al, "Face Customization in a Real-Time digiTV Stream," *Real-Time Imaging VI*, Proceedings of, Vol. 4666, SPIE, 2002.

[12] Peng, C., Lugmayr, A. & Vuorimaa, P., "A digital television navigator", *Multimedia Tools and Applications*, nl. Kluwer Academic Publishers. 17, pp. 121-141. A. Lugmayr, S. Kalli, R. Creutzburg.

[13] Synchronization of MPEG-7 metadata with a broadband MPEG-2 digiTV stream by utilizing a digital broadcast item approach. Real Time Imaging and System Components. SPIE. Seattle,2002.