A. Lugmayr, S. Kalli, R. Creutzburg. 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.

Synchronization of MPEG-7 metadata with a broadband MPEG-2 digiTV stream by utilizing a digital broadcast item approach

Artur Lugmayr[a], Seppo Kalli[a], and Reiner Creutzburg[b]
[a] Institute of Signal Processing, POB. 553, FIN-33101 Tampere, Finland
[b] Digital and Computer Systems Laboratory, POB. 553, FIN-33101 Tampere, Finland

ABSTRACT

Digital, Interaction Television will be deployed in the next few years, where the key-standard for Europe is DVB/MHP[4]. Initial experiences are already reported, but for advanced use scenarios, and application areas basic research issues, such as synchronization of MPEG-7 metadata, retrieved either from the broadcast stream, or the feedback channel are still to be done. In this research work we consider, that metadata will be available from multiple sources, especially commercial emerging MPEG-7 extractors, and streamed to a digiTV client over the broadcast- or feedback network. MPEG-7[1] provides asynchronous or synchronous delivery at transport layer, and DVB/MHP standards provide synchronization via a trigger mechanism. As an example we show a simple hyperlinked television show, where we segmented each frame into elements of interest, to which a hyperlink can be added. The user has the ability to click through those, and obtains additional information about the segmented objects, represented by this region. The key aspects of this paper are as follows:

- Description and introduction of a *Digital Broadcast Item* (DBI) for structuring hyperlinked TV to one entity, that is exchanged between several parties in the broadcast chain based on MPEG-21 standards;
- MPEG-7 metdata extraction models and image segmentation authoring tools for obtaining information about the hyperlinks within a video segment, and within a still frame;
- Metadata transmission and synchronization models;
- MPEG-7 Systems compliant software architecture at client side;

Keywords: MPEG-7, digital television, MHP, DVB, Digital Broadcast Item

1. INTRODUCTION AND OBJECTIVES

A typical digiTV deployment architecture consists of multiple *Broadcast Service Provider (BSPs)*, providing TV content multiplexed into a high-bit-rate *MPEG-2 Transport Stream (MPEG-2 TS)* over terrestrial, satellite or cable infrastructure to the consumer's *Set-Top-Box (STB)*. Optionally the STB is connected either to a home-network solution or directly to the Internet convolving the feedback channel to an *Interaction Service Provider (ISP)* for enabling transactional service types with a *Service Provider (SP)* over wired or wireless solutions. We obtain a unified value chain by involving *Service Editors (SE)*, responsible for the creation and editing of services deployed by the BSP. To state an example, a SP is represented by a travel agency, where a SE creates actual digiTV applications, forwarded to a BSP for broadcast deployment. The consumer might buy a ticket and opens a transactional channel back to the travel agency over an ISP for buying electronic tickets for e.g. a football match.

Multiple parties, therefore various different types of content exchange between them, imply a more contemporary and unified approaches for a more sophisticated and automised solution to configure (e.g. add service applications from a travel agency in form of advertisement add-ons), deliver (either over the feedback- or broadcast channel), interaction facilities (e.g. involvement of transactional services in any form), exchange between providers, track object history,

Further author information: (Send correspondence to Artur Lugmayr)
A.L.: email: lartur@acm.org, Tel.: +358 (0)40 821 0558, Digital Media Institute, Tampere University of Technology, POB. 553, FIN-33101, Tampere, Finland.
R.C.: email: creutzbu@cs.tut.fi, Tel.: +49(0)3381 355 442, Digital and Computer Systems, Tampere University of Technology, POB. 553, FIN-33101 Tampere, Finalnd; and FH Brandenburg, Dept. Computer Science, POB. 2132, D-14737 Brandenburg, Germany.
S.K.: email: seppo.kalli@tut.fi, Tel.: +358 (0)40 053 0267, Digital Media Insitute, Tampere University of Technology, POB. 553, FIN-33101 Tampere, Finland.

enable configuration management, and presentation structure for a more sophisticating digiTV experience. Our efforts within the *future, interaction TV (fiTV)* project (www.futureinteraction.tv) is the convergence of different metadata (commonly known as "data about data") standardization efforts in the field of broadcasting to one unified structure, a *Digital Broadcast Item (DBI).* Within the scope of this research paper, we present *SegmentTV* broadcasted from a BSP, enabling transactional services on the consumer side by simply clicking hyperlinks on video content. We especially point out the DBI approach as convergence solution of different metadata (commonly known as "data about data") exchange standards between key players in the value chain, for configuring and synchronizing content.

The DBI approach is comparable to a simple Web page, represented by the *Hypermedia Tag Markup Language (HTML)* enabling a unified representation, distribution, and presentation of multimedia presentations to the consumer on the Web. In digiTV, currently no structuring elements, of how to align additional content to a broadcast stream exist. To satisfy pure passive users, the main line of presentation must be still one pure audio/video stream without any "service breaks". Especially in a complete digitised environment, this requires high efforts, due to the existence of three different information channel models between BSP, ISP, and the audience:

- **Push Model:** Content distribution relies entirely on a uni-directional path to the participant. Legacy audio/video content is enriched by supportive multimedia information and a local storage device. This feature enables enhanced video-recorder facilities, synchronizing local stored multimedia material to live content, application distribution, and filtering types of personalization. Each service stream is represented by a *Program Elementary Stream (PES)* uniquely identified by an identification number (PID).
- **Feedback Model:** This model enhances the push model by adding a feedback information channel from the participant to the service provider. Besides advanced consumer behaviour driven personalization and monitoring, this model enables full Internet connection, full interactive service types, enriching BSP services by self-sufficient ISP services, download of BSP independent content etc.
- **Distributed Content Model:** Various different service providers (ISPs, BSPs) are part of a distributed broadcast service pool. Services are grouped and structured around multiple broadcast streams. Advanced local content resolution schemes, real-time data exchange (e.g. user profiles), metadata driven distributed services, etc. provide new emerging business models and collaboration of value chain parties.

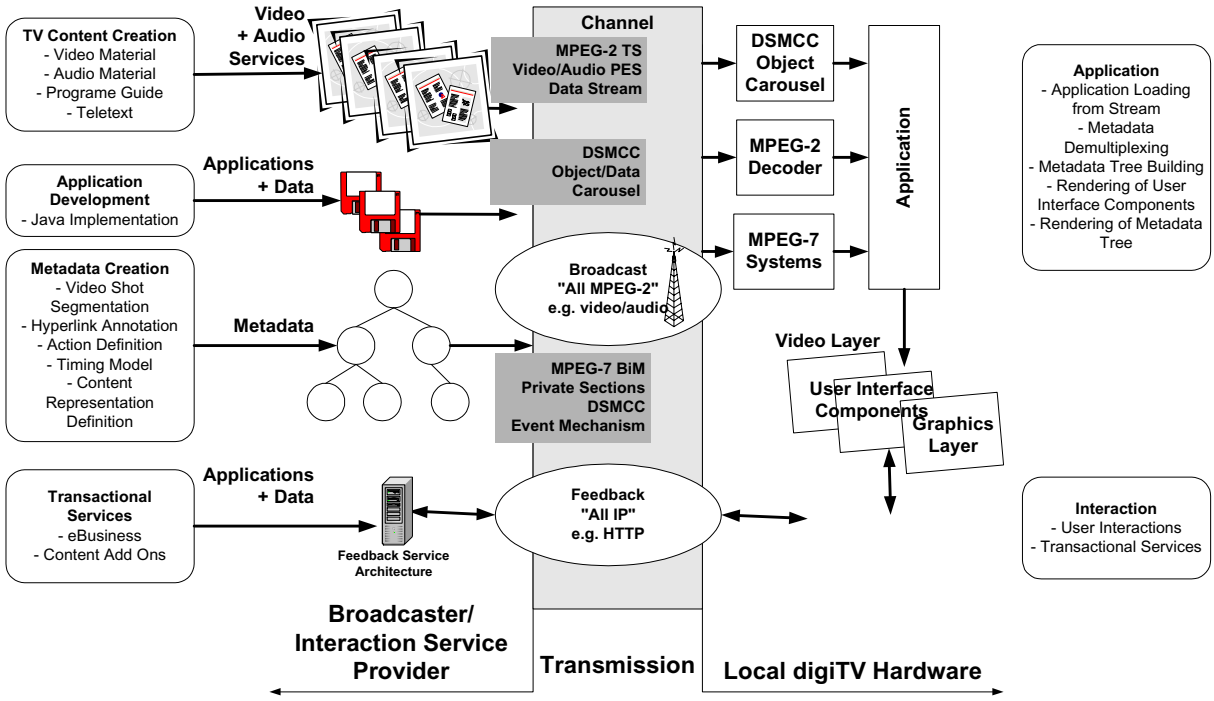

Figure 1. Description of the parties involved in the value added digiTV service chain and their tasks.

Fig. 1 shows a generic overview of the overall system architecture, convolving four levels of creating and adding content. TV content creation convolves mainly commonly known services on digiTV, involving the creational, preoperational, presentational, and representation task of TV content in general (e.g. film shooting, news broadcast creation, *Electronical Program Guides (EPG),* teletext, super-text). Those service types are extended by application development, additional metadata creation schemes, and transactional services. Within the scope of this research work, the other system entities are investigated and presented in further detail.

1.1. Related Research Works

An excellent revision for the exploration for temporal synchronization models has been done by ELISA BERTINO in Ref.[4], "Temporal Synchronization Models for Multimedia Data", describing characteristics, requirements, time-dependencies, temporal relationships, and temporal scenario models/constrains of multimedia systems by pointing to the key research works done within this field. OLIVIER AVARO describes in "MPEG-7 Systems: Overview" the MPEG-7 Systems specification, starting from the general architecture up to the description of the individual MPEG-7 System tools in Ref.[5] The EBU project group, P/META, is defining and representing the information requirements for the exchange of programme content between the high-level business functions of EBU members: Production, Delivery/Broadcast and Archive in Ref.[6], showing a more broadcast related view of metadata. Ref.[7,8] show current projects of how to add hyperlinks to a broadcast show and point out current efforts of web TV. Another metadata related research work has been established by NewsML, providing an XML based standard to represent and manage news throughout its lifecycle including production, interchange, and consumer use in Ref.[10] An excellent overview of software specification techniques for real-time imaging, reviewing unique challenges faced when specifying real-time imaging systems can be found in Ref.[16]

1.2. digiTV related Metadata and Broadcasting Definitions

In the following the key standards and definitions utilized in our system architecture are presented:

- **MPEG-7**, formally named *Multimedia Content Descriptor Interface (MCD)*, aims to create a standard for describing the multimedia content data that will support some degree of interpretation of the information's meaning, which can be passed onto, or accessed by, a device or a computer code[1]. MPEG-7 is developed by the *Motion Picture Experts Group (MPEG)* and standardized by ISO/IEC. Especially for a content driven multimedia environment, as is digital television, it forms a unified and open standard based framework for the exchange of metadata described services in push, pull, and full interactive modes between several parties in the value chain.
- **TV-Anytime** standards, as developed by the TV-Anytime Forum, are on top of MPEG-7 standards focusing especially on TV related matters. Four fundamental objectives were established: exploitation of local persistent storage in consumer devices; network independent delivery of content including various delivery mechanisms (e.g. DVB, ATSC), inter-operable and integrated systems for content creator/provider/consumer, and security structures to protect several parties in the value chain[2].
- **MPEG-21**, also developed by the MPEG group, specifies how to uniquely identify and describe *Digital Items (DI)* and other entities, the relationship between DI and existing identification systems, and the relationship between DIs and relevant description schemes[3]. MPEG-21 is more related to the elements that get exchanged between parties in the value chain and provides structuring elements for resources, its description, and any time of digital content (e.g. electronic tickets in an e-business application).
- **SMPTE**, the *Society of Motion Picture and Television Engineers*, recommends practice of metadata dictionary registry of metadata element descriptions for association with essence or other metadata in Ref.[10] Those standardization efforts focus more on broadcast related issues, and requirements than previously mentioned standards.
- **MHP/DVB,** Europe's emerging digiTV standards are specified by *Digital Video Broadcast (DVB)* standards, where the *Multimedia Home Platform (MHP)* [6] defines a Java based end-user device solution.
- **SUN/JavaTV APIs,** provide an ideal development and deployment platform for emerging digiTV class of interactive services. The Java programming language provides content developers with a high degree of control and flexibility over the "look and feel" of their applications, enabling them to deliver the most dynamic and compelling interactive television experiences to their audience[12].

2. CONVERGENCE OF METADATA STANDARDS TO A DIGITAL BROADCAST ITEM

A DBI definition is based on MPEG-21's metadata definition, where an *Item* groups items and its sub-items, containing *Descriptors* holding information about Items, *Compound* descriptors for binding of a *Resource* to all of its relevant descriptors, and *Container* definitions for grouping Items to logical packages for transport or logical shelves[3]. In the case of a DBI (Container Descriptors) different configurations are available for the creation of SegmentTV.

Fig. 2 shows the configuration of a DBI. In our scenario it is configured with three different sub-items. The *MHP Service Deployment Item* holds all information required to for resolving broadcaster related deployment parameters, such as deployment time, identification numbers, temporal broadcast aspects, required bandwidth, etc. The sources for this metadata are either SP, or BSP related and consists currently of self defined metadata standards, as no general exchange definition exists as far. Information about the content of the broadcast stream and its carrying services itself is included in the *Program Schedule Item*, defining rigid metadata either in form of TV-Anytime's descriptor schemes or rigidly defined *Service Information (SI)*, the common source for every digiTV EPG (please see Ref.[11] for further descriptions). The service, thus the SegmentTV application itself, finds is descriptions within a *Multimedia Presentation Item*, as delivered from a SP and deployed by the use of MHP Service Deployment Item's parameters by the BSP. Its description is based on MPEG-7 multimedia definitions. Another configuration addresses interaction models – the *Interaction Item* – delivered also from the SP for invoking transactional services.

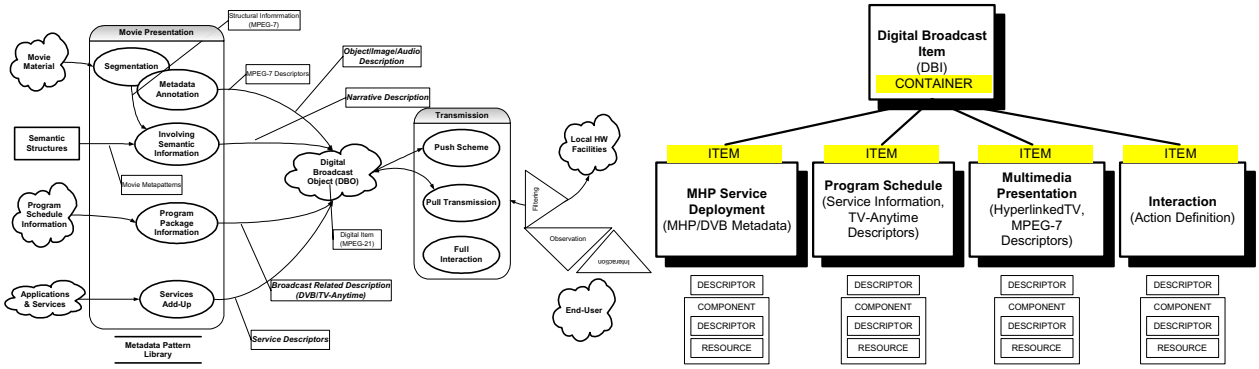

Figure 2. Left: An instantiated Digital Broadcast Item (DBI) - a Digital Broadcast Object (DBO) and its flow process. **Right:** Our configuration of a DBI for the broadcast of a SegmentTV show, consisting of different sub-items.

2.1. The Multimedia Presentation- and Interaction Item Configuration

Three different parts are required to create the configuration for SegmentTV: The descriptor compound is made up by selected MPEG-7 descriptors, defining the representation, presentation, temporal- and spatial relations of hyperlinks within the broadcast show and mostly defined by MPEG-7 descriptors. Secondly, references to resources, such as application and multimedia content files are considered as resource and referenced by an instantiated DBI. Those represent the actual application and its description, furthermore described by the third component within the compound, additional descriptors (e.g. creation history, authors).

Table 1. MPEG-7 defined metadata structures.

Feature	Descriptor/Descriptor Scheme	Description
Video Shots	VideoSegment DS (MDS §11)	Segmented Video Material into shots
Hyperlink representation	StillReagion DS (MDS §11) MovingReagion DS (MDS §11)	Describes the spatial hyperlink region within one or multiple frames
Action Definition	Simple URL (RFC), ImageText(MDS §11), xPath	Action definition
Abstract Segment Description	SegmentDS (MDS §11)	Abstract type of which specialized descriptors are derived

MPEG-7 standard Part 5: *Multimedia Description Schemes (MDS)* describes in §11 tools for describing the structure of multimedia content in time and space for describing segment attributes, structural decompositions, and structural relations, for temporal and spatial segments of multimedia content[1].

- *StillRegion DS*: this descriptor scheme defines two-dimensional spatial regions in an image or frame, either connected or not, and corresponding to an arbitrary set of pixels. Optional the StillRegion DS enables localization and identification of components via the *SpatialLocator* (e.g. position information; component described via a set of vertices), referencing to its containing components (e.g. video sequence defined by a *VisualDescriptorScheme)*, includes timing information (absolute or relative with *MediaTimePoint, TimePoint, RelativeTimePoint* descriptors), and inherits several characteristics from a *Segment DS.*
- *MovingReagion DS*: temporal aspects of a moving region within an image or frame sequence are described with this descriptor. In addition to the parameters described by the StillReagion DS temporal aspects are covered with the *SpatioTemporalLocator* (e.g. movement trajectories), and spatial-temporal decomposition descriptors.
- *ImageText DS:* This tool describes a 2D spatial region of an image or video frame, and is extended from a Segment DS[1]. We utilized this tool for the description of additional hyperlink text.
- *Segment DS:* This tool is an abstract type that describes a fragment or segment of multimedia content. Most specialized Segement DSs are derived from the Segment DSs[1].

2.2. The MHP Service Deployment Item

The MHP Service Deployment Item relates to configuration management of broadcasting equipment, such as:

- **Transport Protocol Parameters for the transmission of User Defined Data:** MHP defines two mechanisms of carrying data to the consumer terminal, Firstly in form of an DSMCC, and secondly as continuous stream, where different protocol types are encapsulated in MPEG-2 private sections. The private section mechanism is comparable to the mechanism of IP packets containing higher protocol types (e.g. eMail) in its payload.
- **Stream Event Temporal and Configurationally Parameters:** Two different versions of stream messages, one for carrying the stream objects not related to the DSM-CC stream events, and one including a stream carrying the DSM-CC stream events are defined in Ref.[13] and named as *Normal Play Time (NPT)* reference descriptors with a time base associated to a stream or event object. As defined in Ref.[13] three use scenarios are applicable: do-it-now events without NPT's scheduling mechanism; single continuous time base for enabling a single continuing interactive production at the BSP; and unique time base for each program requiring the suspension of global time bases[13].
- **Content location and resolution:** A DVBLocator[13] `dvb:// <original network ID>.<transport stream ID>.<service ID>.[.<event ID>] [;<component tag>]` uniquely identifies services deployed on a DVB compliant digiTV architecture. The locator is comparable with a URI, localizing web content on the Internet. Both, original network ID and transport stream ID , where the service ID points to concrete content assets, deployed in the broadcast schedule of one broadcaster. Events ID and component tags are used for identifying continuous events.
- **Digital Storage Media, Command and Control (DSM-CC) Configuration:** Transmitting asynchronously, and repeatedly at a certain frequency a file system like structure to the consumer's STB is the purpose of this mechanism. The configuration of it is defined by multiple parameters, such as service ID (unique service identification), provider- and service name, carousel ID (mechanism for transmission of data within MPEG-2 TSs), *Application Identification Table (AIT)* (table defining the location of applications in the MPEG-2 stream), stream PIDs, event description for controlling the application life-cycle, file path definitions, required bandwidth etc.

2.3. The Program Schedule Item

More related to the description of the present, following upcoming program content are firstly rigidly defined metadata as based on SI, and TV-Anytimes programme description. Both metadata types have are presented with a Navigator software to the end-user and are multiplexed within a TS and must be demultiplexed in the STB before the Navigator can use it. The SI adds information that enables STBs to tune automatically to particular services and allows services to be grouped into categories with relevant schedule information[11]. Where SI defines simple structures and are kept on a minimal required information level, TV-Anytime's metadata definitions are a rich on additional content descriptions (e.g. description of actors, film processing). A more detailed description about DVB-SI can be found in Ref.[11] We use

both description types only for announcing content, and its services (e.g. HyperlinkTV application included, currently TV programme).

2.4. Temporal Synchronization and Transmission Models for DBI Related Services

Within this section we describe briefly how a metadata tree can be transmitted and synchronized at the client-side, referring to Ref.[5], describing exhaustively MPEG-7's terminal architecture and *Binary format for Metadata (BiM)* for transmitting binarized metadata structures. Ref.[4] provides an overview and points out solutions of synchronizing the metadata tree inter-, and intra stream. Each metadata description file has a tree form that can be transmitted, verified, and accessed via a flow of *Access Units* (either binary encoded or textual) containing D, DSs, and DLLs. At client side the stream is parsed to reconstruct the MPEG-7 description tree[5]. Each tree part can be updated independently with an update, add, or delete command.

Table 2. Temporal synchronization models.

	Entity of Operation	**Example**	**Characteristics**
Media Layer	Single continuous metadata stream	PES BiM Application Stream	Device independent interface of operations en-/decapsulation processes
Stream Layer	Group of media streams Single Media Stream	MPEG-2 DVB stream A whole program inc. applications and content	Interstream synchronization Intrastream synchronization
Object Layer	Temporal synchron. Specification as input	Multimedia presentation Hyperlinked TV including object information	Intra-/Inter stream synchronisation based on lower layer calls
Transaction Layer	Higher layer Internet protocols (HTTP, etc.)	XML based eBusiness	Asynchronous / Synchronous
Human Layer	Human interactions	Activation of Hyperlinked TV	Asynchronous
Material Preparation Layer	Preoperational entities	Extraction of metadata	Soft deadlines

Temporal synchronization consists of a media layer, which represents the lowest level in the hierarchy, an application operates directly on a single continuous media stream which is seen as composed by a sequence of information units. At the stream layer an application operates on a group of media streams as well as on a single stream, whereas the information units within each media stream are not visible. At object layer takes a synchronization specification as input and produces a presentation that meets the temporal constrains in the specification[4]. We extended this model by the transaction layer, involving user-inputs and other feedback types in companion with the human layer. The material preparation layer addresses more material preparation issues, such as video segmentation, and mostly has to meet soft deadlines, due to a more available time span.

3. SEGMENTED TELEVISION

Segmented television is one DBI, consisting of the previously described parts. It covers several synchronization layers, but the most relevant is the object layer, where MPEG-7 descriptors are utilized to describe temporal-, and spatial aspects of hyperlinks, representing kickable objects during the broadcast show.

3.1. Material Preparation

Material preparation tasks address video segmentation, DBI configuration, and service implementation. Firstly, we segmented the video broadcast stream into shots, by utilizing the research work of M. Höynck[14], and his IENT-Cut automatic video segmentation software, to obtaining MPEG-7's video segment description schemes. For each shot for duration of longer than 20 seconds (to provide set-up times at the STB, and the invoking of user transactions) we selected a key frame and annotated hyperlinks and their actions to it.

3.2. Deploying Hyperlinked Television

For our sample scenario "Hyperlinked TV" we utilized our broadcast laboratory facilities, consisting of a commonly used digiTV deployment infrastructure, transmitting a MPEG-2 TS at a bit-rate of 6.780.562 b/sec, feedback architecture, and DVB complaint set-top box. The core element for our test-trials was a superior real-time MPEG-2 TS multiplexer with multiple input sources for data, video, and audio broadcasting. In combination with a DVB compliant setup-box and a satellite modulator we were able to transmit a DVB compliant digiTV stream over a satellite network, carrying our own simple television program including applications and metadata structures. Five different services, uniquely identified by a DVBLocator (e.g. `dvb://1.1.1001` for the accessing of our hyperlinked application identified with the service ID 0x1001, or a video broadcast identified with service ID of 0x10 with a locator pointing at `dvb://1.1.10`), convolved our broadcast show.

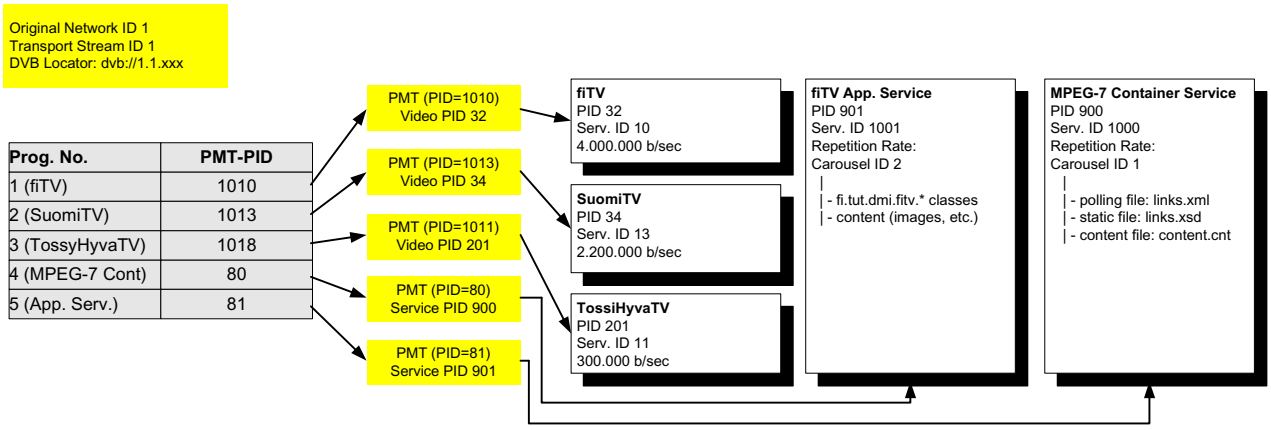

Figure 3. Deployment configuration.

The television programme comprised three TV channels (fiTV, SuomiTV, and TossiHyvaTV) for audio/video PES transmission at a bit-rate of 6.500.000 b/sec together, one DSMCC object carousel for application carriage (100.000 b/sec), and three different facilities for carrying metadata including still-frames forward to the client side and synchronize there locally with audio/video content. Current DVB compliant hardware development efforts do not support frame capture form a live stream. Our workaround is based on the segmentation of video material in shots (VideoSegment DS) with a minimum duration of approx. 2 minutes to provide adequate client setup- and end-user interaction times. For each shot a representative frame or video scene, annotated with MPEG-7 (StillRegion-, MovingRegion DS) metadata, and relating action definitions is decoded at the client side. The end-user gets a brief notification, either if he wants to activate the hyperlink mode for content browsing or remain watching TV. Fig. 3 shows our complete test configuration including its parameterizations.

3.3. MPEG-7 BiM Transmission Modes

Our implementation significantly varies from introduced methodologies, as our local MHP compliant hardware did not support all features required to support all transport facilities. In the following facilities for the carriage of our DBI are presented:

- *DSMCC Object Carousel*: At the broadcaster side one object carousel (in our case identified by `dvb://1.1.1001`, carousel ID 2) is facilitated for the transmission of metadata and related hyperlinked material (e.g. still frame) for the currently running shot. The carousel contains a constantly multiplexed DS file for verification purposes at client side, constantly polled content file (either still frame or video scene), and a constantly polled MPEG-7 BiM file describing the content file. A real-time, soft deadline constrained file update mechanism at broadcaster side realizes the update of the polled files for each video shot delivered as television programme. The client is notified of changes with the DSMCC event mechanism. This architecture is limited in BiM streaming facilities to simple file updates, allows shot accurate hyperlink interaction, allows fast client setup times, and is simple to realize.

- *Private Section Mechanism*: Both, MPEG-7 BiM and shot representative content, are encapsulated within a private section stream within a MPEG-2 TS. This procedure enables continuous streaming, therefore frame accurate metadata annotation, but requires longer client setup times, is more complex to realize, consumes higher client performance, and requires hard deadline constrained transmission.
- *NPT Events*: Each NPT is associated with a service (e.g. `dvb://1.1.10.100`) and contains private data, containing MPEG-7 BiM's structures including shot related content. The use of NPT events for hyperlinked TV is rather limited due to its signalling type of use, but enables more accurate scheduling of the related material. The utilization of this mechanism only for SegmentTV is not desirable, but enhances the previously mentioned methods by additional facilties: timebase simulated event monitoring, scheduling of Hyperlink TV announcements, timebase synchronization, and carriage of MPEG-7's time base descriptors for enhanced synchronization purposes of two transmitted files.

Fig. 4 shows the broadcast related flow process, including its mechanisms.

3.4. Client Software Architecture

The client architecture follows a component based approach, where each component communicates via events. We strictly kept the *Model View Controller (MVC)* approach to separate each underlying task. Our application parts can be categorized in following modules:

- **MHP Compliant Components:** MHP compliant component modules access, control, and coordinate MPEG-2 TS and its functionality. In our experiments we utilized a *TVStreamVisualization* and *TVTuning* component for controlling and tuning video/audio input, and rendering it on a STB. A *DSMCCComponent* controlled the access modes to MPEG-2 DCM-CC facilities, and delegated action commands to the application layer, from which further processes are initialised.
- **Application Component:** The central element for interconnecting different components, invoking transactional services and user inputs, controlling application flow, and performing their life-cycle is done by the application.
- **XML Components:** XML components perform the task of MPEG-7 System and parse several XML related content files related to a DBI.

The flow process at client side is presented in Fig. 4. The use of the DSM-CC component is two-folded; firstly it is used for application loading, secondly for the delivery of a virtual file system to the STB client containing the BiM metadata structures, implicating the use of two object carousels for this purpose. The XML Component performs several XML parsing related tasks, convolving MPEG-7 System's BiM decoding. For saving bandwidth we considered a lookup table implementation, for textual resolving XML tags. The metadata tree is continuously updated and rendered in the hyperlink mode of the application. The actual shot frame snapshot, therefore segmented still frames including hyperlink information, is transmitted within the same carousel, as BiM structures. Besides the metadata mechanisms TV functionality is present in the current software implementation. During the broadcast show the user can select if he would like to experience hyperlinks upon which the application changes to hyperlinked mode, where objects in frame snapshots can be selected, and a transmission channel opened for further processing.

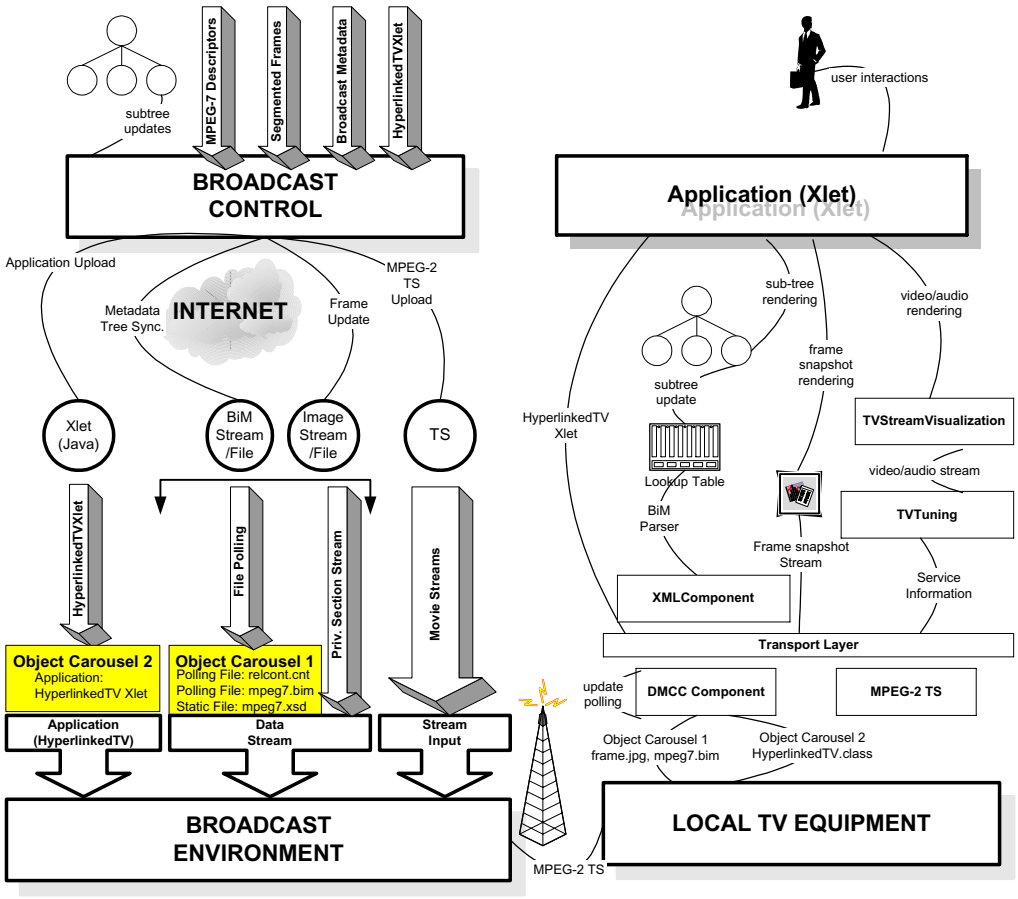

Figure 4. Broadcast-, and client architecture flow diagram.

4. EXPERIMENTS AND CONCLUSION

4.1. Client Performance

We tested our application on a MHP compliant TriMedia STB TM1300[15] running at 160 MHz, 4 Mbytes flash memory, 4 Gbyte hard disc, Personal Java based Java VM, MPEG-2 hardware decoder, etc. We can distinguish between application-, and continuous metadata update performance values. For both we evaluated memory-, and time consumption. For the first we obtained a rather long *application set-up times* for application signalling and loading from approx. 4.3 sec. Memory usage was 2.982.192 bytes free memory of 4.194.296 after complete application loading with several components. *Rendering times* for both types were between 21ms and 27ms for painting user-interface, and graphical components. Continuous *metadata update performance* addresses DSM-CC polling and event delegation, requiring less than 0.5 sec. More problematic were XML parsing processes, consuming 2.6 sec. for parsing a metadata update file containing in average 22 elements (approx. 14ms for each) including attributes. 33 ms where required for destroying our application. For a continuous metadata file (approx. 80 elements plus attributes), holding basic DBI information of original size 2.66 KBs, we achieved a compression rate of 29% (778 bytes) after compression and binarization. Binarization alone was not enough through MPEG-7 AUs header overhead blew the file to 4.88 KBs. Involving a lookup-table mechanism compressed the metadata file down to 653 bytes, and provided slightly faster parsing times (approx. 1 sec.).

Figure 5. Screenshots of our sample application (from left to right). Navigator software during an "Aquarium" video broadcast; Segmented TV for browsing through hyperlinks and its action defined by a small image "Fishing is fun"; Yellow active hyperlink and red passive hyperlink.

4.2. Broadcasting Performance

Figure 6. Our test-laboratory (from left to right). TV facilities including a couch for "Couch Potatoes"; MHP compliant STB; Broadcasting architecture.

At BSP side we performed our tests with three test streams, of a few minutes video material. In the material preparation phase we semi-automatically segmented and annotated video material to obtain metadata descriptions. Missing data was purely manually configured. We performed three carousel updates per video stream, sufficient for our test environment. Besides already described data rates, each application was multiplexed at an AIT rate of approx. 0.8/10. Kbits/s, configured for auto-start, at input rates of approx. 100/360 Kbits/s. Metadata file updating was done automatically, by a file polling mechanism, and uploading new files over the web. The whole stream multiplex of 188 byte packets consisted of 41% video/audio data, 2% application related data, and 0.38% SI, and the rest were stuffing or NULL packets.

4.3. Conclusions

The overall result was a sophisticating solution for merging the paradigm of hyperlinks and digiTV together to one entity. Most problematic are client performance issues, due to too high metadata parsing times. The end-user experiences a rather long response time. But the approach of utilizing MPEG-7 metadata descriptors instead of rigidly defined MPEG-4 standards for converging live and virtual objects provided a more than sophisticating solution for our application. In our future work we highly focus on a more optimised decoding, and try to experiment with other protocol types as soon as the relevant hardware is available. On our Web pages (www.futureinteraction.tv) you can find some of our results of done research works.

ACKNOWLEDGEMENTS

We would like to thank all the project team members of the fiTV project at the Tampere University of Technology, Hypermedia Laboratory at the Tampere University, and the Helsinki University of Technology. We thank them for all their help, discussions, and research materials. Special credits go to Olli Stroem, and Pasi Rustanius for their fishing tips and indirect test material provision.

REFERENCES

1. ISO/IEC JTC 1/SC 29/WG 11, *MPEG-7 Standard and Recommendations*, Pattaya, December, 2001.
2. TV-Anytime Forum, *TV-Anytime Standards and Recommendations*, http://www.tv-anytime.org.
3. ISO/IEC JTC 1/SC 29/ WG 11, *MPEG-21, Information Technology – Multimedia Framework – Part 3: Digital Item Identification and Description*. Sidney, July, 2001.
4. E. Bertino, and E. Ferrari, "Temporal Synchronization Models for Multimedia Data," *IEEE Transactions on knowledge and data engineering*, Vol. 10, No. 4, July/August, 1998.
5. O. Avaro, and P. Salembier, "MPEG-7 Systems: Overview," *IEEE Transactions on circuits and systems for video technology*, Vol. 11, No. 6, June, 2001.
6. R. Hopper, *P/META: Metadata Exchange Standards*, EBU Project Group, EBU Technical Review, September 2000.
7. P. Hoschka, *Television and the Web*, http://www.w3.org/Talks/1999/0512-tvweb-www8/Overview.html, W3C. 1999.
8. V. M. Bove, et. al, "Hyperlinked television research at the MIT Media Laboratory", *IBM Systems Journal,* Vol 39, No. 3&4, 2000.
9. International Press Telecommunications Council (IPTN), NewsML, http://www.newsml.org/NewsMLweb/webpage.xml.
10. SMPTE. SMPTE standards and recommendations. www.smpte.org.
11. C. Peng, A. Lugmayr, and P. Vuorimaa, A Digital Television Navigator, *Journal of Multimedia Tools and Applications*, Vol. 17, pp. 121-141, Kluwer Academic Publishers, 2002.
12. SUN, Java Technology in digital TV, Vers. 1.0, November 2000, http://java.sun.com/products/javatv/index.html.
13. Digital Video Broadcast (DVB), *MHP: Multimedia Home Platform Specification,* Vers. 1.0.2 Rev. 2, European Broadcasting Union (ETSI,EBU), Februar, 2002.
14. M. Höynck, C. Mayer, J.-R. Ohm, "Application of MPEG-7 Descriptors for Temporal Video Segmentation", *Proceedings of the SPIE International Conference on Electronic Imaging - Storage and Retrieval for Media Databases*, Vol. 4676, pp. 347-358, San Jose (CA), USA , January 2002
15. Philips, *TriMedia Documentation*, September 2001. www.philips.com.
16. P. Laplante, and C. Neill, "An overview of software specification techniques for real-time imaging," *Real-Time Imaging VI*, Vol. 4666, pp. 18-29, Proceedings of SPIE. San Jose, January, 2002.