Multimedia broadcast and internet satellite system design and user trial results

J. Cosmas, T. Itagaki, K. Krishnapillai and A. Lucas

Brunel University, Uxbridge, Middlesex UB8 3PH, UK
E-mail: john.cosmas@brunel.ac.uk E-mail: T.Itagaki@brunel.ac.uk
E-mail: Kannan.Krishnapillai@brunel.ac.uk
E-mail: alan.lucas@brunel.ac.uk

Abstract: The EU funded project, System for Advanced Multimedia Broadcast and IT Services (SAMBITS), has created an enhanced and synchronised, multimedia terminal for merging satellite broadcast and internet telecommunication services in a way that efficiently combines the large bandwidth of the broadcast channel and the interactivity of the internet. This paper proposes a novel broadcast and internet service concept, illustrates this concept with two service scenarios and develops a system architecture to demonstrate the range of key benefits provided by these new technologies. It then describes the interactive multimedia terminal that was used for consuming this new service concept. Finally, the results of the user trials on the terminal are presented and discussed.

Keywords: DVB-S; MPEG; MHP; satellite broadcast and internet service.

Reference to this paper should be made as follows: Cosmas, J., Itagaki, T., Krishnapillai, K. and Lucas, A. (2005) 'Multimedia broadcast and internet satellite system design and user trial results', *Int. J. Services and Standards*, Vol. 1, No. 3, pp.336–357.

Biographical notes: Prof. John Cosmas obtained a BEng honours degree in Electronic Engineering at Liverpool University in 1978 and a PhD in Image Processing and Pattern Recognition at Imperial College in 1987. In 1983, he joined Imperial College as a Research Student, in 1986 Queen Mary & Westfield College as a Lecturer in Digital Systems Design and Telecommunications, in 1999 Brunel University as a Reader in Multimedia Systems and the subsequently as a Professor of Multimedia Systems in 2002. His research interests are design and experimentation of new 4th generation TV and telecommunications services and networks, multimedia content and databases and video/image processing. He has been or is contributing towards eight EEC research projects throughout his career and has published over 60 papers in refereed proceedings and journals. He is the Director of Multimedia Systems research within Brunel IT (BIT) Laboratory.

Dr. Takebumi Itagaki obtained a BEng degree in Applied Physics at Waseda University in Tokyo (Japan), a Postgraduate Diploma in Music Information Technology at City University in London and a PhD at University of Durham with a thesis 'Real-time Sound Synthesis on a Multi-processor Platform' covered issues on DSP, distributed and parallel processing, and real-time sound synthesis. In 1998 he was appointed as a Lecturer at Glasgow Caledonian University, thence in 2000 as a lecturer at Brunel University where he has been involved with EU-IST RTD Projects in multimedia and digital TV; 3D

MURALE, SAMBITS, CISMUNDUS, SAVANT, CONFLUENT and INSTINCT.

Mr. Alan Lucas obtained a BEng honours degree in Computer Science at University of Glasgow in 2000 and has been persuing PhD research into terminal application architectures for converged broadcast and telecommunication services.

Mr. Kannan Krishnapillai obtained a BEng honours degree in Electronic Engineering at Brunel University in 2000 and has been persuing PhD research into content and service description tools for converged broadcast and telecommunication services.

1 Introduction

The original concept of broadcast television was essentially a passive technology with a linear presentation of audio-visual media. The modern digital broadcast television introduces a number of new concepts, namely interactive, enhanced and synchronised television.

Interactive television can have two different types of interactivity, namely: local and global interaction (Chen, Feng and Yu, 2000). With local interaction, users interact with the television whilst with global interaction the television interacts with the service provider's network to extend the interactivity of the terminal to outside the boundaries of the home. The concept of local interaction is not new within broadcast industry but the amount of influence or degree of interactivity has now increased. Users can now influence not just channels but focus their influence within individual television programmes. Global interactivity, however, is a new concept to broadcasters, while being fundamental to the internet. It is this concept that has led to a convergence of broadcast and internet services, supported by the terminal.

Enhanced television is not a new concept that belongs to the digital television generation (Thomas, 2001). With analogue television, broadcasters discovered that although video and audio required the majority of the analogue signal's bandwidth, there was still enough bandwidth (although very limited) to include other forms of data such as subtitles and signers. Enhanced television consists of additional information/data that is multiplexed with the traditional audio and video. This additional information/data allows the user to view/use additional content which is associated with the traditional television programme. Within the strictest sense of the word 'interact', then one could argue that enhanced television is also interactive television. However unlike interactive television, no return channel is required i.e. no communication is required between the user and the broadcaster/internet service provider because all additional information/data can be sent via the broadcast channel and therefore only local interactivity is involved. However, with the introduction of digital television, the true potential of enhanced television can be realised. One of the great advantages digital television has over analogue television is the ability to compress digital multimedia, thus allowing greater amount of information to be contained within the broadcast signal (transport stream). In the latter part of the 20^{th} century, the Moving Pictures Experts Group (MPEG), which is a working group of the International Standards Organisation/International Electrotechnical Commission

(ISO/IEC), started a series of open multimedia compression techniques that has produced the MPEG-1, MPEG-2, MPEG-4 and MPEG-7 standards. MPEG-1 was concerned with the audio visual coding for storage media, MPEG-2 with audio visual coding for transmission, MPEG-4 with audio video coding of arbitrary shaped images and 3D graphics and MPEG-7 with the annotation of multimedia content.

Synchronised television provides additional content to the main MPEG-2 AV programme which is dependant on the time line of the programme. This can be achieved by using MPEG-7 meta-data to describe the additional content including timing information. This meta-data can be used to determine when the additional content could be available to the user.

The advent of interactive, enhanced and synchronised television will require content producers to rethink how to design their programmes so that they can be easily managed and readily and comfortably accessed using these new media. This implies that the rules that govern TV and internet programme making and journalism for these new services will need to adapt in order to capture and maintain the interest of viewers whilst not continuously distracting their focus of interest. An experimental terminal system architecture was conceived to test new concepts for future multimedia broadcast and internet services so that user trials can be carried out. These trials were required to determine the acceptability of additional, interactive multimedia content with different categories of TV broadcast programmes such as documentary and entertainment programmes. With the purpose of conducting these user trials, a standards compliant architecture of an experimental future multimedia broadcast and internet terminal was designed and built. This architecture of the full system consisted of the broadcast programme production, service provisioning and the multimedia terminal systems.

2 System concept

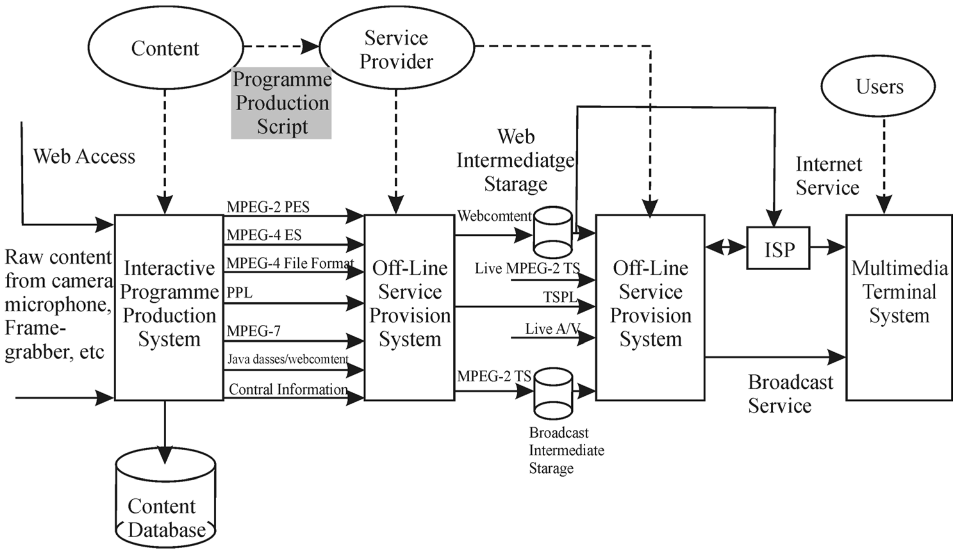
The multimedia broadcast and internet system consists of an Interactive Programme Production Sub-system (IPPS), Off-Line Service Provision Sub-System (Off-line SPS), the Online Service Provision Sub-System (Online SPS) and the Multimedia Terminal, as shown in Figure 1.

A satellite transmission system was chosen for digitally broadcasting the service since these were the earliest digital broadcast systems to have been deployed and the most likely to be the first to introduce innovations to their multimedia services. The US American Television Standards Committee (ATSC) system has mainly been applied to cable and terrestrial broadcasts whilst the Japanese Integrated Services Digital Broadcasting (ISDB) system has mainly been applied to terrestrial broadcasts. The DVB-S system was chosen because this system has experienced the greatest acceptance from satellite broadcasters in the world and is likely to be the most widely deployed.

The IPPS was designed for use by content producers for the iterative creation of content for broadcast and internet services. It consists of a suite of software tools available at the Personal Computer (PC) desktop for the capture, creation, encoding, indexing/integrating and visualisation of content. This service architecture was required to enhance the audio/visual broadcast TV programmes with additional audio/visual clips, 3D graphics, internet pages and a database and to efficiently compress these different types of media for transmission using open source standards. Since the DVB standard uses MPEG-2 TV audio/visual encoding and transmission, it was decided to use the

ISO/IEC suite of standards for encoding arbitrary shaped video (MPEG-4) (Pereira, 2000), 3D graphics (MPEG-4) (Kauff *et. Al*, 1998) and databases (MPEG-7) (Pearmain *et al.*, 2002; Chang, Sikora and Puri, 2001; vanBeek, 2000) while using the internet's Hypertext Markup Language (HTML) from World Wide Web Consortium (W3C). The new multimedia services were designed containing MPEG-2, 4, 7 and internet elements that allow for both local interaction to locally stored content and remote interactivity to content stored on servers.

Figure 1 SAMBITS system



The main reason for developing an off-line SPS was that the MPEG-4 technology is not mature enough for commercial DVB multiplexing equipment to handle properly. There is no MPEG-2 real-time multiplexer on the market today that is able to compliantly insert MPEG-4 elements in MPEG-2 transport streams. The online SPS implements as far as possible the features/capabilities of a real-time broadcast and internet service provision platforms, including the possibility to modify components of the broadcast service shortly before or even during play-out of the service and to allow for dynamic updates of content in the service.

The main functionality of the terminal was to decode and display MPEG-2 A/V and MPEG-4 A/V streams, locally stored MPEG-4 AV and MPEG-4 3D files and the outcome of searches on MPEG-7 files. Browsing on the internet is also required. The choice of software architecture for the terminal consisted of OpenTV (Hirtz and Zwing, 1996), Multimedia Home Platform (MHP)[1] (Evain. 1998), Multimedia and Hypermedia Experts Group (MHEG) (Colaitis and Jourdain, 1995; Kraft and Stolp, 1998; Meyer-Boudnik and Effelsberg, 1995) systems. OpenTV is a closed standard for developing interactive TV applications on a set-top box and has extended its middleware to support DVB-MHP. MHEG is a content format standard that allows multimedia content to be presented within the context of a television environment and can be represented as a set of files. The MHP provides the whole set of technologies that are necessary to implement digital interactive multimedia in the home. It was felt that the choice of terminal system architecture was a critical issue for testing the concept of additional multimedia content within different types of TV broadcast programmes. The MHP system was chosen because it is an interactive, open source standard that has been widely adopted by broadcasters in the world and is thus becoming a good candidate for the world-wide standard for the interchange of services.

3 Service concept

A service concept that goes beyond the linear presentation of audio/visual TV programmes was required to be conceived. One of the main methods for embellishing new services is through the application of novel, service discovery mechanisms. Traditional methods of service discovery have been based on search tools for the internet or Electronic Programme Guides (EPG) for the TV broadcast programming. Contemporary service discovery methods have been based on an event of some type. The use of message events has been a way service agents have notified users of an event that a user has expressed an interest in through subscription (Bais *et al.*, 2002). For example, when a goal in a football match has been scored then the user is sent a message notifying him of the event. The use of location-based events has been a way mobile communication services have notified users of content that is related to an attraction at a physical location through subscription (Giordano, Chan and Habal, 1999). When a user roams to a location sufficiently close to an attraction, she is notified of available multimedia content that is related to it. The notion of programme events relates multimedia content to a new segment of video. When a video programme experiences a segment change, the broadcast network pushes multimedia content to the terminal that is related to that segment of video. The user is then free to access this additional multimedia content through local interaction during the segment of video. Whilst watching this additional content the main audio/visual programme can be stored on local disk for time-delayed viewing at a time when the viewing of additional content has ended.

New multimedia services were created based on programme and message event principles. Figure 2 shows the interactive multimedia broadcast service concept. The main broadcast audio/video programme is streamed as a series of programme segments together with the arbitrary shaped video of a sign language presenter. A set of related content is repetitively transmitted with the main programme during the period of a programme segment. This consists of the annotation of the main programme, annotation of programme segments, additional multimedia and HTML content and the annotation of multimedia content. All this annotation and related content are repetitively transmitted during the segment period because the user may start to view a programme at any time during the programme's transmission. Message events are transmitted asynchronously delivering textual content of a real time event that may have occurred.

The MPEG-2 transport stream (TS) provides several mechanisms to multiplex content, as shown in Figure 3. It has defined five different profiles/applications for data transport: data piping, data streaming, multiprotocol encapsulation (MPE), data carousels and data services that require periodic and object carousels. The MPEG-2 main programme AV and the MPEG-4 sign language 2D presenter content was multiplexed and transmitted as two DVB, Packetised Elementary Streams (PES). Non-real time multimedia content was broadcast in the DVB carousel providing background information. This consisted of the description of the main programme segment and the programme segment's related content was transmitted in the DSM-CC object carousel sections. Message events are transmitted over the DSM-CC private data multi-protocol Internet Protocol (IP) encapsulator section.

Figure 2 Interactive multimedia broadcast service concept

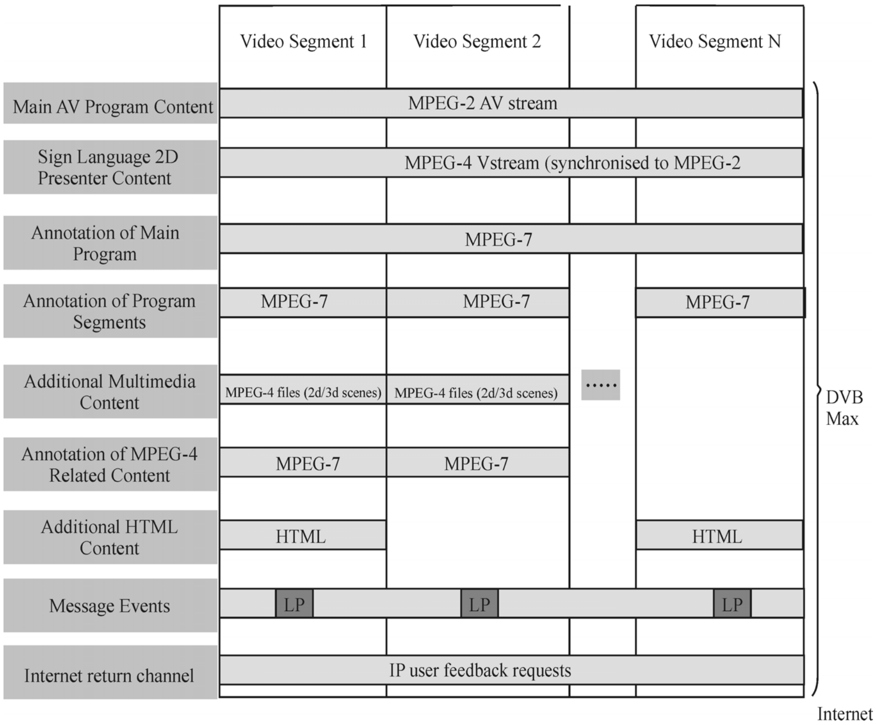
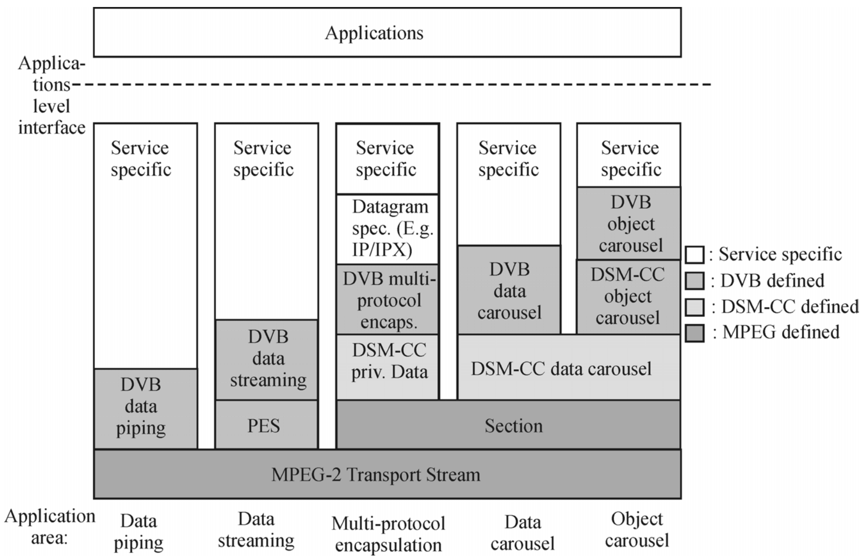


Figure 3 DVB profiles/applications for data transport



4 Service scenarios

Two service scenarios were conceived, namely: ‘The Ballad of Big Al’, an educational documentary programme; ‘Eurovision Song Contest’, a light entertainment show. These two programmes were selected because their styles were quite different with each making different demands on the attention and concentration of the viewer.

4.1 Dinosaurs

‘The Ballad of Big Al’ is a ‘Walking with Dinosaurs’ special programme made by the BBC. The main components of the programme are listed in Table 1. A short version of this was used as the main AV programme. The four background movie clips were taken from ‘Big Al Uncovered’, which describes the science behind ‘The Ballad of Big Al’. Introductory web pages were taken from the BBC’s website for broadcast transmission to the terminal. However, the viewer was able to web-browse outside the ‘walled garden’ of web pages broadcasted to the terminal, offering a much richer experience and stimulating the use of telecom network for accessing additional high value content. For example dinosaur games could be accessed through the BBC’s web pages. The importance of the narrative in the programme highlighted the usefulness to viewers with auditory handicaps of having a sign language interpretation. This could be watched by viewers with auditory handicaps without distracting other viewers.

Table 1 Media components of the ‘Walking with Dinosaurs’ programme

	Description	*Data Type*	*Media Type*	*Multiplexor*
Main programme	Main A/V content	MPEG-2 A/V	Stream	PES
	Programme description	MPEG-7 file	File	Data carousel
	Sign language	MPEG-4 video	Stream	PES
Related / background material	Four Background movies	MPEG-4 A/V	File	Data carousel
	Sign language related to background movies	MPEG-4 video	File	Data carousel
		Mux together with movies to a scene (BIFS)		
	Info on BBC web-page	HTML page	File	Data Carousel
		HTML link	Link	Component of MPEG-7 description
	Games on BBC WEB-page www.bbc.co.uk/dinosaurs	HTML link	Link	Component of MPEG-7 description

The viewer could watch the programme as normal (Shot1: Normal TV watching). If the viewer suffered from auditory handicaps, and their user profile was set accordingly, the viewer could choose to watch the programme with sign language interpretation (Shot 2: Main programme with Signing). When some background information became available, this was indicated to the viewer. They could choose to ignore this, continue watching the main programme and perhaps view it later by going to the programme guide, or they could view the background content instead of the main programme. In this example, the background information might be some video clips (Shot 3: Background video clips without signing). As with the main programme, the viewer could choose to turn on the signing. Other background information included web pages (Shot 4: A webpage related to the programme). If the viewer required further information, at any time during the programme, they were able to search for it using the in-built search engine (Shot 5: Search querying). The search results may be web pages or video clips, either of which may have been previously recorded by the terminal or on the internet. Programme Info is also available throughout the programme. This gives a summary of the whole programme, and also of each scene (Shot 6: Programme Info). It shows the additional information available.

4.2 Eurovision song contest

The Eurovision Song Contest involves a number of contestants from European countries and Israel taking part in performing a song and then each country votes on which song they considered was the best. Since 2000, the European Broadcasting Union has also been web casting the Eurovision Song Contest on the internet. This means, that the interested viewer can also watch associated content, for example A/V clips of former winners and streams from backstage cameras. However, both web casts up to now have not worked for most viewers, because of the incredible interconnectivity of hundreds of thousands of viewers at the same time. The internet is not ready yet to serve such a large number of viewers with the same content at the same time. The proposed broadcast and internet system provides a perfect solution to such an event with the capability to stream additional content over the DVB channel. This means that the clips of former winners and the backstage camera can be encapsulated in the DVB transport stream during the event, when lots of people want to access this information at the same time. The same clips can be kept available on the internet later, when only a few people want to have access to this information. The main components of the programme are listed in Table 2.

Table 2 Media components of the Eurovision Song Contest programme

	Description	*Data Type*	*Media Type*	*Multiplexor*
Main programme	Main A/V content	MPEG-2 A/V	Stream	PES
	Programme description	MPEG-7 file	File	Data carousel
Related material	3D models / scenes for interactive navigation	MPEG-4 3D scene	File	Data carousel
	Past contestants' movie clips	MPEG-4 A/V	File	Data carousel
	Backstage Camera	MPEG-4 video	File	Data carousel

The main AV programme consisted of the Norwegian entry 'On My Own' by Haldor Laegrid followed by the Greek entry 'Die for You' by the Antique group. The backstage camera movie clips were taken from actual backstage shots during the performance. Additional movie clips of past contestants were taken from past television archive material. Introductory web pages were taken from the EBU's website for broadcast transmission to the terminal. Again the viewer was able to web-browse outside the 'walled garden' of web pages transmitted to the terminal. A 3D hall of fame was constructed where the viewer was able to navigate within this 3D world and select portraits of past winners. When the winner was selected then a surround sound audio clip of the selected past winner was played. The attitude of the viewer within the navigation tool dictated the direction of audio that the user experiences.

The viewer could watch the main show as normal (Shot 7: Eurovision Song Contest) or could choose to watch what was happening backstage (Shot 8: Backstage camera). The audio and visual information of the back-stage was transmitted through a DVB-channel and allowed its access to millions of people at the same time as the main broadcast. In addition to the programme overview and background video clips and web pages, this example also showed an interactive 3D scene (Shot 9: MPEG-4 3D Scene). This showed ‘Hall of Fame’ of the Eurovision Song Contest’s previous winners. The viewer can navigate in a 3D virtual hall and by selecting the pictures of the winners whereupon an extract from their winning entry is played. This example also showed the use of news flashes (Shot 10: A News Flash).

5 Multimedia terminal system

The terminal was based on a conventional set-top box (STB) from Fujitsu-Siemens which has already been introduced into the market[2] and its main components are shown in Figure 4. Its functionality is limited to decode through an MPEG-2 demultiplexor and display MPEG-2 A/V streams as well browsing the internet to some extent. The manufacturer specific Application Programmer’s Interface (API) for navigation and control lets the broadcaster provide services only to a limited number of boxes.

The terminal was enhanced to access and decode multiple types of media streams in parallel, which came from different sources (multiple DVB PESs, DVB object carousels, DVB User Datagram Protocol (UDP) IP encapsulation, local disc, internet, etc.) as shown in Figure 6. This system was extended by adding MPEG-4 and MPEG-7 functionalities and used DVB-MHP from IRT, Germany (Brefort, 2001) for the application software which provided all necessary functions of an STB. MHP used Windows as the underlying Operating System; however, the MHP concept can easily be ported to other OS, namely Linux (Hogan, 2000). Significant functionality to MHP was extended, mainly in two areas: storage manager and user interface. As the content may be displayed only upon request, it was necessary to locally store the broadcasted content. The storage management had to deal with a large variety of related data and with different delivery channels: DVB, via the internet, or local storage. The user interface was based on the MHP navigator and had to deal with multiple player instances and with complex relationships between data. In particular, the user interface provided a homogeneous look-and-feel interface to all player components within the broadcaster’s ‘walled garden’ which was unfortunately not realised in full because MPEG-4 provides its own navigation controls. Internally, the user interface interfaced to all other components. User interactions included giving a feedback to broadcasters via an internet connection. Such an internet link also allowed easily integrating internet navigation as well as locating additional content, e.g. MPEG-4 files, MPEG-7 files, HTML pages. The protocol stacks that were needed to connect to the internet were typically part of the underlying Operating System.

Several other components were developed and integrated (Figure 6). One key element was an MPEG-4 decoder audio and visual streams and files. The MPEG-2 main programme AV and the MPEG-4 sign language 2D presenter content was multiplexed and transmitted as two DVB, Packetised Elementary Streams (PES). The streamed

MPEG-4 content was synchronous to the main programme. Non-real time multimedia content was broadcast in the DVB carousel providing background information. This concept enabled the user to decide whether to view this content or not. However, the user was also able to retrieve MPEG-4 content from the internet and display it simultaneously. Interactive 3D scenes, e.g. virtual rooms, were broadcast and the user was able to navigate through the scenes independently from the main programme. Hence, the terminal integrated a streaming enabled 3D capable MPEG-4 player.

Figure 4 System outline of a conventional set-top box

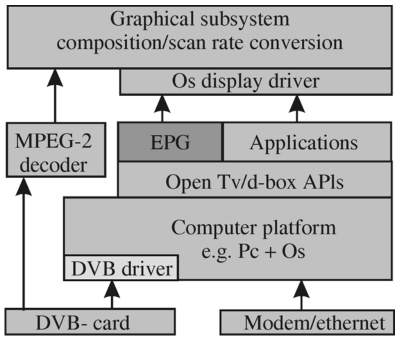


Figure 5 System outline of SAMBITS multimedia consumer terminal

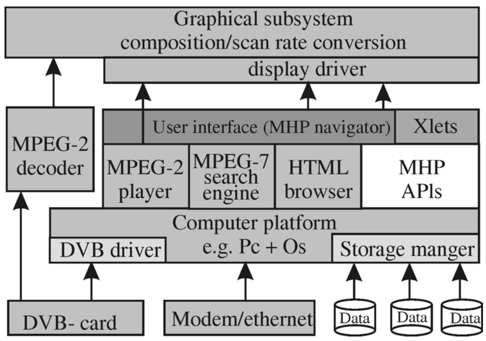
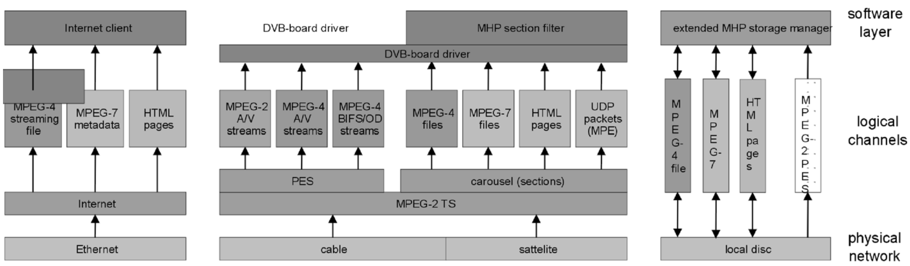


Figure 6 Overview of the data types and input channels handled by the SAMBITS platform



Another key element was an MPEG-7 parser for timeline management and an MPEG-7 search and query engine. MPEG-7 meta-information was inserted as additional information into the transport stream, complemented the video content and facilitated users to retrieve background information (content search). Moreover, MPEG-7 descriptors enabled the location of specific sequence sections in the current programme or in material stored locally (Ileperuma, Lalmas and Rollicke, 2001). The type of information was much more elaborate than what could have been realised with DVB-SI tables, as the descriptors included but were not limited to textual descriptors. Also, an MPEG- 7 description was associated typically to each media stream like MPEG-2 or MPEG-4, such that there may be multiple MPEG-7 metadata files broadcast in parallel. The terminal system used the MPEG-7 description of the main programme to locate related material and signal the availability to the consumer. Before the descriptors were parsed either for time line management or in the search engine they were first decoded in an MPEG-7 binary decoder.

Furthermore, the concept integrated an HTML- browser for internet browsing, but also as a front-end to the search engine. Since multiple video (MPEG-2, MPEG-4) and graphical windows (MPEG-7, HTML, EPG, and user interface) were combined, stacking the windows required sophisticated alpha-blending so that all relevant information was visible at the same time. Although the system lacked hardware support for 3D graphics, the processor was powerful enough to perform 3D rendering in software for the scenarios.

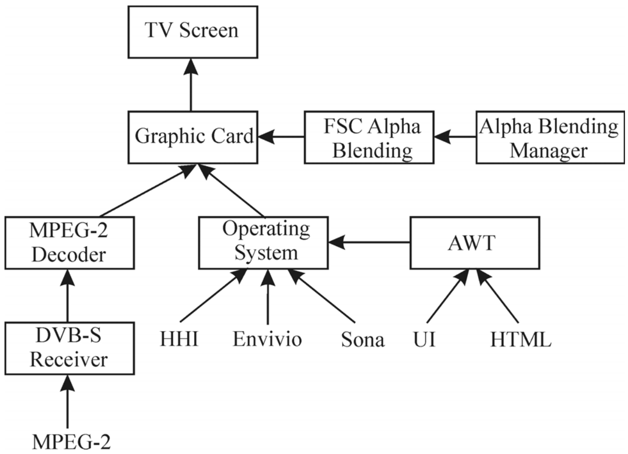
5.1 Alpha blending architecture

As in any commercial set top box, the graphic output is a usual TV screen. Therefore the graphic card was specially designed to avoid flickering when dealing with such output. In the terminal, the graphic composition played an important role because of the different components which needed to be harmoniously placed on the screen. The composition used three different graphic sources that overlap and were handled through the Alpha Blending Manager (ABM).

The MPEG-2 stream represented the first graphic source and contained the main programme. Once the DVB-S receiver had been tuned, the MPEG-2 stream was sent to the decoder which placed the video on the screen. The size and position of the background video was fixed to full screen and to level zero. All other components were blended into this component at level one as shown in Figure 7.

The Java and Windows graphic output shared the same level one and were treated in the same way from the ABM. It would have been a nice feature if the use of Windows graphics system could be removed to allow portability to other operating systems, but the use of different Windows-based players did not make this possible. The operating system was responsible for placing them correctly on the screen, but the ABM decided which regions were required to be visible or their level of transparency. When the terminal was switched on, the user interface was initialised and the ABM created a region at level one (foreground) that matched the screen location of the user interface rendered by AWT and set the correct alpha value for it. Each time the user interface changed its layout the ABM modified the regions to suit the new distribution. The same process took place for the HTML player.

Figure 7 Graphics layers 0 and 1



The rendering of the other components (Song player, HHI decoder and Envivio player) were very similar. They avoided going through Java AWT and used the facilities of the operating system, which provided higher performance. The most complete situation occurred when one of these components was on the screen. Then, the background video, the user interface (player control) through AWT and the player itself through Windows were rendered at the same time, and the ABM set up a region for each of them.

In the cases of the 3D MPEG-4 Player and the MPRG-4 video decoder, Magic Colour was used. The procedure was the same but the motivation was different. The objective in the 3D MPEG-4 Player component was to provide a transparent background of the 3D scene keeping at the same time the opacity of the 3D object. Since the background was designed to be homogeneously black, its value was set as Magic Colour. When using the MPEG-4 Video Decoder in combination with arbitrary shaped objects, Magic Colour was used. The MPEG-4 Video Decoder supported output frames that contain alpha information, but its format was not appropriate for a fast rendering. So, the alpha information was avoided and homogeneously black background which surrounded the 2D object was used. Table 3 lists the alpha values assigned to each component.

Table 3 Alpha values for system components

Component	*Alpha Value*	
	Description	*Value / Magic Colour*
MPEG-2 video	Opaque	–
User interface	Semi-transparent	180
HTML player	Opaque	255
HHI decoder	Opaque with magic colour	255 / 0
Song player	Opaque with magic colour	255 / 0
Envivio player	Opaque	255

5.2 Storage manager (SM)

Compared to the rudimentary Storage Manager (SM) of MHP, the terminal SM had to accomplish a rather complex task. The storage manager accessed different data delivery channels, such as using MHP APIs: DVB PES; DVB carousel; DVB private sections; DVB IP sections; internet socket (HTML, streaming). In order to locate the streams, the DVB tables were accessed and processed. The SM had to handle many types of media, in parallel, as shown in Figure 6 and had to handle many different files in a timely manner.

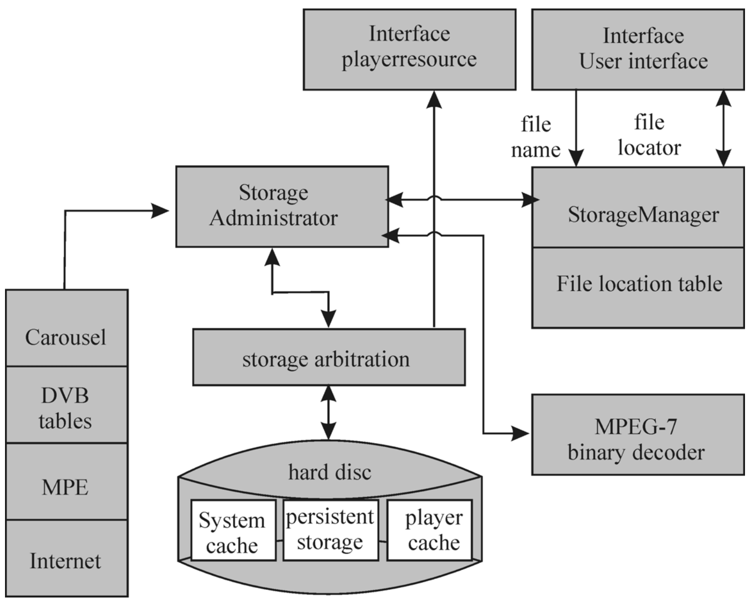
At the beginning of a programme the SM located the carousel, downloaded the MPEG-7 file for the MPEG-2 A/V stream and decoded the file using the MPEG-7 binary decoder. Parsing the file revealed related material that was, in case of non-real-time, content downloaded from the carousel and stored locally. For streaming content, the SM identified the relevant section IDs or PIDS, tuned in, and started downloading and passing the data to the relevant players. As this required the player to be ready, this was handled in close co-operation with the user interface.

All incoming MPEG-7 data was redirected also to the MPEG-7 search engine, which created an internal database for searching later. The main MPEG-7 files were also passed to the user interface for timeline management.

As some type of content was related to the main programme, in particular background video clips, the content of the carousel was dynamically switched. The SM had to identify a new carousel and download the necessary files. In order to keep the system load at a minimum, a permanent scanning of DVB tables was avoided by signalling new carousels through UDP over DVB. All locally stored files could be located for later play-out. The SM controller was responsible for content management, file insertion, file deletion and maintaining content description lists. Moreover, the SM controller handled all data access requests, for an example Xlet requests to store broadcasted MPEG-4 files or player requests to access a locally stored MPEG-4 file.

Technically, the Storage manager subsystem was partitioned into a manager and an administrator, as shown in Figure 8. The manager instructed the administrator to save data, e.g. from the DVB carousel to persistent storage. The appropriate filename was created by the administrator and returned to the manager.

Figure 8 Storage manager architecture



5.3 User interface

The User Interface (UI) was designed such that it could be operated with a standard remote control, needing just five buttons: Info, Overview, Channel, Search and Extras. It implemented interfaces to all players and also integrated an Espial HTML browser.

Figure 9 shows the UI control flow tree containing the button that was presented to the user and Figure 10 shows the buttons presented to the user at the top level of the UI. The Info displayed text information regarding the current video segment that is being presented to the user. This information consisted of a segment heading and brief introduction text. The user was notified of the change of segment when the Info button flashed red. The Overview button displayed a list of all segments of the current television programme. In addition, the list showed what additional content was associated with what segment (even though the user may not be able to access the content if the content is not available, i.e. not yet downloaded) and provided an abstract textual description for each segment. The Channel button displayed all available television services. The available television services were displayed as a list. The Search button displayed a set of keywords in a table format and was used to return additional content, if that content was available at the current time period. Once the search had been completed the results (additional content) of the search was displayed as a ranked list. A red indicator bar was displayed next to each entry showing how close the match is. The Extras button displayed a list of currently available additional content for the current time period i.e. current segment.

Figure 9 User interface flow tree

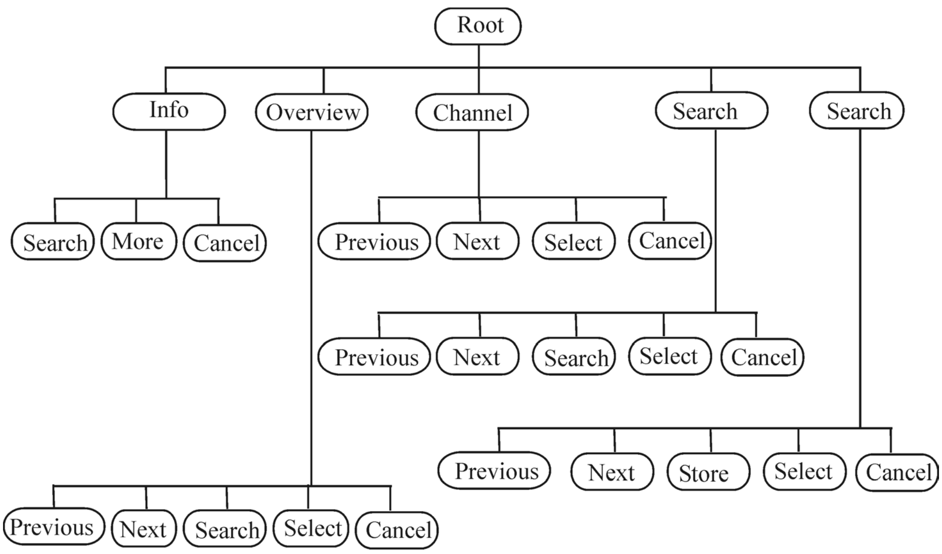
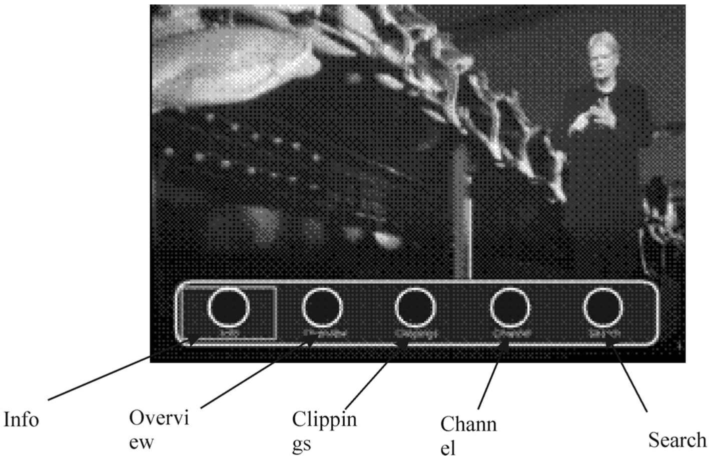
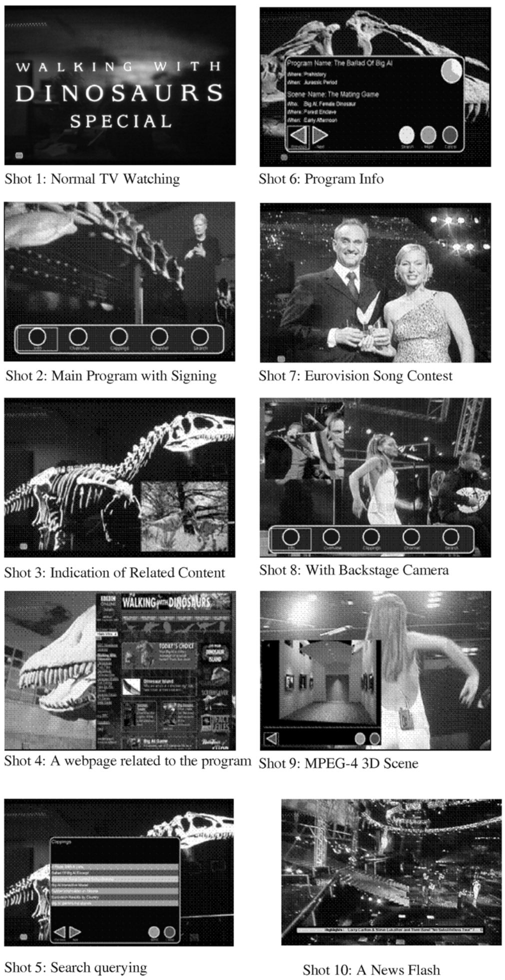


Figure 10 Main UI buttons presented to the user





'Previous', 'Next' and 'Cancel' buttons were used to navigate to previous and next elements in lists related to overview information, channels, searches engine and extra information and to navigate back to top level of the user interface tree, respectively. The 'Search' and 'Select' buttons within the 'Overview' branch were used to initiate searches and highlight additional content, respectively. The 'Search' and 'More' buttons within the 'Info' branch were used to activate search engine screen and scroll to next page of introduction text, respectively. The 'Select' button within the 'Channel' branch tunes into a service within the transport stream, depending on the currently highlighted element in the list of channels. The 'Search' and 'Select' buttons within the 'Search' branch were used to initiate the MPEG-7 search engine within a subset of keywords and select keywords for the search, respectively. The 'Search' and 'Select' buttons within the 'Extras' branch were used to persistently store highlighted content to hard disk and display highlighted additional content, respectively.

The user interface controlled the signalling of related material by parsing the MPEG-7 file of the main programme and tracking the segment timeline described in this file. If the MPEG-7 file signalled related material by means of an URL, the UI checked with the SM whether this file was already available, and after a positive feedback signalled the availability of related content to the user. If the user decided to view that content, the UI launched the required player and provided the file locator to the player. Moreover, the UI acted together with the HTML browser as a front-end to the MPEG-7 search engine. The user interface also handled user and terminal preferences that were described by MPEG-7 files.

5.4 MPEG-4 players

The additional audio/video content was encoded employing the core profile of MPEG-4 in order to enable arbitrary shaped objects, while audio was encoded in the AAC format. Since the demonstration scenarios also encompassed interactive navigation through 3D scenes, a 3D capable MPEG-4 player was integrated as well. However, as the player development was not part of the project, the player module was obtained from external sources. Since no one MPEG-4 player had all the required functionalities, three different MPEG-4 players were used. For decoding of 2D A/V scenes the terminal employed the ENVIVIO player. The HHI player was used for the integration of an arbitrary shaped signed language presenter into the MPEG-4 scenes where the user could switch the content on/off as required. For 3D scenes, the player from the European Union project 'Portals for the Next Generation' (SONG) was integrated.

The players were activated from the user interface that also passed the file locator. The player then retrieved the content either from a file or from a streaming device. The player controls for AV-content were handled through the user interface while the 3D-navigation remained within the player.

5.5 MPEG-7 engine

For data search and retrieval, an MPEG-7 engine was provided. Moreover, the engine was used to locate specific content being broadcast. For that reason, the MPEG-7 engine connected to the MPEG-7 related streams and processed them immediately. The result was stored in an internal database.

Users were now provided with the opportunity to search for related material. Therefore, based on the MPEG-7 descriptors in the current broadcast, queries were predefined and signalled through the user interface to the user. If selected, the search engine scanned the internal database and searched for related material. The ranked search results, which were an XML file, were rendered through the HTML browser.

5.6 HTML browser

The HTML browser was the common element to access internet content. As the STBs were designed to operate mainly in the 'walled garden' of broadcasters, all content was designed such that the terminal required no specific functions from the browser. As MHP is JAVA based, the Espial ESCAPE Java browser was used.

6 Terminal user trials

6.1 User trial procedure

The user trials were conducted under laboratory conditions. Users were directed to read a 'User Instructions' guide and then asked to carry out a set of tasks on the terminal. Observations of the users were made during the laboratory experiment including the users' mistakes, verbal comments and questions. Finally the users completed a questionnaire and discussed the system. The results of the terminal user trials were plotted on Graphs 1.1–1.25.

6.2 Evaluation of user trials

On the whole users found it easy (Graph 1.25: Average = 1.9/5.0; Graph 1.24: Average = 2.2/5.0; Graph 1.23: Average = 1.8/5.0) to learn how to operate the system and the exploration of new features was also relatively easy (Graph 1.22: Average = 1.8/5.0; Graph 1.21: Average = 2.1/5.0;) using trial and error. Remembering specific rules to enter commands was slightly more difficult (Graph 1.20: Average = 2.6/5.0). While users found steps to complete a task logical (Graph 1.18: Average = 2.2/5.0) and successfully completed them (Graph 1.17: Average 2.5/5.0), they found the number of steps per task too numerous (Graph 1.19: Average 3.0/5.0). On the whole, users felt that the sequence (Graph 1.15: Average = 2.4/5.0) and layout of menu screens were acceptable (Graph 1.13: Average = 1.8/5.0; Graph 1.12: Average = 2.1/5.0). The system response time (Graph 1.10: Average = 2.9/5.0) and stability (Graph 1.8: Average = 2.2/5.0) were acceptable to users although noticeable. Users found the general impression of the system to be acceptable, although they found it easier to use (Graph 1.4: Average = 1.9/5.0) and more powerful (Graph 1.5: Average 2.1/5.0) than the other categories. The commercial prospects of such a system was thought to be very high (Graph 1.1: Average = 1.8/5.0).

6.3 Analysis of additional user comments

When asked about the usefulness of additional content while watching the primary television programme, it was revealed that the effectiveness of the additional content was dependent on the nature of the primary television programme. Users believed that programmes such as documentaries, political and news programmes as well as light entertainment would be able to take the greatest advantage from having additional content while for movies, dramas and programmes that have a dynamic and complicated plot, the additional content might serve simply to distract the users from following the plot.

For the additional content to be of use, users indicated that the synchronisation between the primary television programme timeline and the additional content was very important. The users were uncertain about the amount of 'useful' information that the additional content contained.

Users found that the keyboard controls for 3D MPEG-4 player were too sensitive for users to navigate 3-D scenes and in addition they wanted all controls to be combined into one device e.g. mapped to the remote control.

With regard to video screens, users wanted to resize window controls, to move the video and HTML screens to suit their own needs and to be provided with an indication of loading times for MPEG-4 clips and 3D MPEG-4 Player. Users found the fact that there was no auto start of MPEG-4 clips annoying.

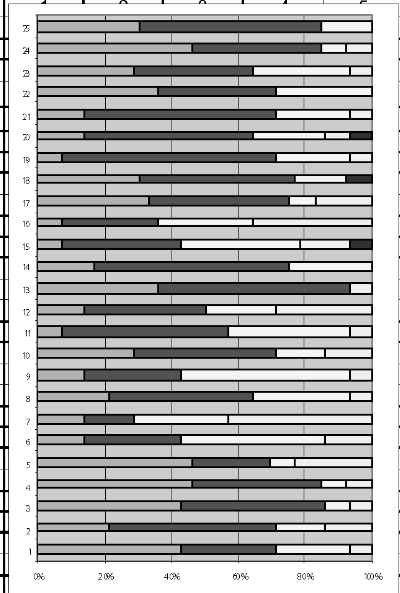
With regard to the User Interface Control Panel, users wanted the 'INFO' button to be next to the 'EXTRAS' button. This was due to the amount of interaction between these two buttons. Users wanted the 'Channel' button to be displayed first (most left) on the user interface. Users wanted a wrapping function on the User Interface Control Panel, thus providing a 'short-cut' from the left most button to the right most button and vice versa. Many users wanted the User Interface Control buttons to be mapped onto the remote control as 'Hot Buttons' rather than being displayed onto the screen.

With regard to the User Interface Menus, users wanted the use of the 'Up and Down' arrow keys to navigate between items in the menus and wanted the 'Select and Cancel' functions to be mapped to the remote control as 'Hot buttons' rather than being displayed on the screen. Menu transparency caused items within menus to be difficult to read because of the background video. It was not clear to users when items within the menus were highlighted and when menu control buttons were highlighted.

With regard to the Search Engine Controls, the layout of the 'Keywords' made it difficult for the user to navigate. Navigation controls were too difficult. In particular, most users expressed a wish to use the 'Up and Down' arrow key in conjunction with the 'Left to Right' arrow keys on the remote control to select 'Keywords'. The number of operations required to perform a search was considered too much. Users wanted the 'Select and Cancel' functions mapped to the remote control rather than being displayed as buttons on the screen. Users reported that the search engine screen overlay intruded too much over the main video. They wanted a smaller search engine screen, and the ability to resize and move the search window. The 'Keywords' terms were not self-explanatory and some users did not understand the concept of information association with the 'Keywords'.

Graph 1 Results of user trial

			Average		
25	Learning	Learning to operate the system: Getting Started	1.9	Easy	Difficult
24		Learning to operate the system: Learning advanced features	2.2	Easy	Difficult
23		Time to learn to use the system	1.8	Fast	Slow
22		Exploration of new features by trial and error: Exploration of new features	1.8	Easy	Difficult
21		Exploration of new features by trial and error: Discovering new features	2.1	Easy	Difficult
20		Remembering specific rules of entering commands	2.6	Easy	Difficult
19		Tasks performed in a straight-forward manner: Number of steps per task	3.0	Just right	too many
18		Tasks performed in a straight-forward manner: Steps follow logical sequence	2.2	always	never
17		Tasks performed in a straight-forward manner: Completion of sequence of task	2.5	clear	unclear
16	Screen	Design of User Interface	2.1	clear	confusing
15		Sequence of menu screens: Next screen in a sequence	2.4	predictable	unpredictable
14		Sequence of menu screens: Beginning, middle, end of task	2.6	clear	confusing
13		Layout of menu screens: Amount of information on the menu screens	1.8	adequate	inadequate
12		Layout of menu screens: Arrangement of information on the menu screens	2.1	logical	illogical
11		Layout of menu screens: Design of the menu screens (colors, fonts, etc)	2.8	helpful	irritating
10	System	System speed / response time for operations	2.9	good	too slow
9		Stability: Operations are	2.1	dependable	unsdpendable
8		Stability: System failures occur	2.2	Seldom	often
7	General Impression	Wonderful	2.3	Wonderful	Terrible
6		Satisfying	2.4	Satisfying	Frustrating
5		Stimulating	2.2	Stimulating	Dull
4		Easy	1.9	Easy	Difficult
3		Powerful	2.1	Powerful	Not Powerful
2	System Prospects	This application enhances TV experience	1.8	enhancement	No enhancement
1		System will be used in the future	1.8	yes	no



During the design of the user interface, it was assumed that the remote control would provide simpler controls to the application if the number of functions that were mapped on to it were limited. However, following the user trials it was discovered that users wanted more functions mapped to the remote control, thus reducing the number of operations required to execute the features of the application. Although the terminal was not a commercial product, from these user trials the following factors should be implemented for such a system to become a successful commercial product in the future:

- the price of the Set Top Box must be competitive with similar products in the market
- the extra features such as the presentation of additional content must be sufficient for the customer to purchase the system
- the user interface controls must be 'easy' for all users to navigate between the Set Top Box functions
- the user interface on the Set Top Box must be more pleasing.

7 Conclusions

This paper proposes a model system architecture to demonstrate future multimedia broadcast consumer terminals and outlines some specific details of the production and service provisioning systems used to create and deliver the broadcast and internet programme as well as the demonstrator terminal. Example scenarios have been defined that demonstrate the range of key benefits provided by these new multimedia broadcast technologies. The terminal integrated 2D and 3D capable MPEG-4 players, an MPEG-7 search engine, and an HTML browser which enabled a significantly enhanced broadcast experience. In order to integrate all these components, the system development led to various extensions of the current MHP standard in terms of the storage management, the MHP navigator (user interface), and connectivity to the internet. The terminal concept enabled a broad range of applications ranging from MPEG-4 streaming, adding additional A/V content, 3D navigation to search for related content by means of MPEG-7. The connection to the internet provided access to personalised content, additional material, and a return channel to the broadcaster. The demonstrator exhibited the huge potential of enhancing broadcast programmes in combination with internet. Since the technology is still in the development phase, all possible functionalities have not yet been integrated. Howeve, the terminal concept seems to be flexible enough to allow for and to support many more features.

User trials were conducted and reported on the terminal. Users were generally enthusiastic about the concept of the proposed new services but suggested the need for significant improvements in the use of video screen and in the design of the user interface. Users believed that the effectiveness of the additional content was dependent on the nature of the primary television programme where programmes such as documentaries, political and news programmes and light entertainment would be able to take the greatest advantage from having additional content, while for movies, dramas and programmes that have a dynamic and complicated plot, the additional content serves simply to distract the users from following the plot. The outcome of this preliminary research into interactive multimedia broadcast and internet programmes has shown that the underlying journalistic principles may be quite complex and it may take a number of years for journalists to fully understand this type of new media because of the complex interaction of many different types of media, despite the fact that internet journalists have adapted to their new media quickly.

8 Acknowledgements

The authors gratefully acknowledge the support for this work which was funded by the EU under the IST programme in the project SAMBITS (System for Advanced Multimedia Broadcast and IT Services), IST-1999-12065.

References

Bais, M., and Cosmas J. et al (2002) 'Customized television: standards compliant advanced digital television', *IEEE Transactions on Broadcasting*, June, Vol. 48, No. 2, pp.151–157.

Brefort, A. (2001) 'Multimedia home platform (MHP) standard for digital television', *Fernseh-und Kino-Technik*, November, Publisher: Huthig, Germany, Vol.55, No.11, pp.678–680.

Chang, S., Sikora, T. and Puri, A. (2001) 'Overview of the MPEG-7 standard', *IEEE Transactions on Circuits and Systems for Video Technology*, June, Vol. 11, No. 6, pp.688–695.

Chen, Y.Q., Feng, Y.L. and Yu, S.Y. (2000) 'Digital TV set top box and its development toward interactivity', *Shanghai Jiaotong Daxue Xuebao*, February, Vol. 34, No. 2, pp.286–290.

Colaitis, F. and Jourdain, P., 'MHEG and its profile for ITV applications (Interactive TV)', *IEE Colloquium on Interactive Television*, London, UK, (Ref No. 1995/159) IEE 1995, pp3/1–8.

Evain, J.P. (1998) 'The multimedia home platform-an overview', *EBU Technical Review*, Publisher: Eur. Broadcasting Union, Switzerland, Spring, No. 275, pp.4–10.

Giordano, A., Chan, M. and Habal, H. (1995) 'A novel location-based service and architecture', *Sixth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, PIMRC'95. Wireless: Merging onto the Information Superhighway (Cat. No.95TH8135). IEEE. Part Vol. 2, pp.853–857, New York, NY, USA.

Hirtz, G. and Zwing, R. (1996) 'OpenTV: operating system for interactive television', *Fernseh-und Kino-Technik*, March, Vol. 50, No. 3, pp84–89.

Hogan, B. (2000) 'Embedded Linux-poised for growth?' *Embedded System Engineering*, February-March, Vol. 8, No. 2, pp.62–63.

Ileperuma, D., Lalmas, M. and Rollicke, T. (2001) 'MPEG-7 for an integrated access to broadcast and intemet data', *Proceedings of International Conference on Media Future*, Florence, Italy.

Kauff, P., Ohm, J.R., Rauthenberg, S. and Sikora, T. (1998) 'The MPEG-4 standard and its applications in virtual 3D environments', *Proceedings of 42nd Asilomar Conference*, Monterey.

Kraft, A. and Stolp, R. (1998) 'The MHEG-5 standard for interactive digital TV services', *Fernseh-und Kino-Technik*, December, Huthig, Germany, Vol. 52, No. 12, pp.714–716, 718–721.

Meyer-Boudnik, T. and Effelsberg, W. (1995) 'MHEG explained', *IEEE Multimedia*, Spring, Vol. 2, No. 1, pp26–38.

Pearmain, A., Lalmas, M., Moutogianni, E., Papworth, D., Healey, P. and Rölleke, T. (2001) 'Using MPEG-7 at the consumer terminal in broadcasting', *Proceedings WIAMIS 2001 Workshop on Image Analysis for Multimedia Services*, Tampere, Finland, May 16 to 17, 2001, ISBN 952–15–0625–3, pp.165–170.

Pereira, F. (2000) 'Tutorial issue on MPEG-4 standard', *Signal Processing: Image Communication*, Special Issue on MPEG-4, Elsevier Press, Vol. 15, No. 4.

Thomas, G. (2001) 'Serving up data for enhanced DTV programmes', *SMPTE Journal-Society of Motion Picture & Television Engineers*, May, Vol. 110, No. 5, pp.299–303.

vanBeek, P., Benitez, A.B., Heuer, J., Martinez, J., Salembier, P., Smith, J. and Walker, T. (2000) 'MPEG-7 multimedia description schemes XM', July, Bejing, ISO/IEC JTC 1/SC 29/WG 11/N3464.

Notes

1 ETSI TS 101 812, (2000) Digital Video Broadcasting (DVB), 'Multimedia Home Platform (MHP) Specification 1.0', Sophia Antipolis, 2000.

2 ITU-T Recommendation H.222.0 | ISO/IEC 13818–1: (1994) Information technology – 'Generic coding of moving pictures and associated audio information: systems', Geneva, 1995.