UNIVERSAL MULTIMEDIA ACCESS FROM WIRED AND WIRELESS SYSTEMS*

Andrew Perkis,[1] Yousri Abdeljaoued,[2]
Charilaos Christopoulos,[3] Touradj Ebrahimi,[2]
and Joe F. Chicharo[4]

Abstract. Personal computing and communication devices such as computers, personal digital assistants (PDAs), and mobile phones are moving to their next generation in which the end user will be able to access a multitude of information with a single device either locally or through a network. One likely trend in future personal computing and personal communication is that there will not be a single but several equivalent devices available to users allowing access to information in various forms. Each user, depending on his/her needs would access one or several among them depending on the situation and his/her preference. Using existing protocol mechanisms, in this case, a mapping and negotiation of resources during connection setup would be performed, which would remain in place throughout the life of the connection.

This paper provides an overview of universal multimedia access (UMA), a concept for accessing multimedia content through a variety of possible schemes, and discusses some of the issues that arise regarding its deployment. In particular, UMA will provide a solution for adapting the delivered content when users attempt to access their choice irrespective of their terminal characteristics and communication infrastructure, as opposed to the assumption that the content remains fixed and the objective is to deliver the original content at all times. This recognition represents the impetus for the development of media descriptions and hence UMA; that is, the notion that valuable information can be derived from a variety of conversions of a multimedia content source.

The issues discussed are future requirements on content servers and multimedia viewers, media conversions, UMA protocols, and UMA network architectures. The problems addressed are quality of service issues in network solutions for multimedia communications

and reconfigurable architectures and network control based on source adaptations through media conversions and transcoding.

Key words: Multimedia communications, media description, transcoding.

1. Introduction

The state of the art in distribution of rich multimedia content has been brought forward by the deployment of the Internet and digital TV. The Internet carries all kinds of audiovisual information in a wealth of formats and rates. Standardization efforts of the International Standards Organization (ISO) have given us the JPEG family for still image coding, and the Moving Pictures Experts Group (MPEG) family for audiovisual coding and manipulation in digital TV services. The International Telecommunications Union (ITU) and other independent bodies bring similar telecommunications standards forward.

The MPEG family of standards deals with the representation, compression, and transport (MPEG-1, MPEG-2) as well as the manipulation (MPEG-4) of audiovisual information. MPEG-7 [24] is to deal with the semantics-based representation of content, which will enable management of multimedia in a broad sense. The main objective of the MPEG-7 standard is to include additional information to describe multimedia content. Such additional information is usually referred to as meta data. In this paper, we use the term "descriptions" to avoid any confusion with different definitions of meta data used in other contexts. The description schemes are defined using a special language called the Description Definition Language (DDL), also part of the standard. The latter is an extension of the Extensible Markup Language (XML).

Universal multimedia access (UMA) [26] deals with the delivery of images, video, audio, and multimedia content under different network conditions, user and publisher preferences, and capabilities of terminal devices. A major motivation behind UMA is to enable terminals with limited communication, processing, storage, and display capabilities to access rich multimedia content. UMA presents a solution for wired and wireless clients to access the same content server, each receiving content enabled for their client's capabilities. UMA is an important class of applications considered under the MPEG-7 framework, dealing with, among other topics, media conversion and linking MPEG-7 to real-life applications, such as multimedia communications over a large range of platforms and access technologies. This again will ensure that content generators are not faced with generating various versions of the same content and its synchronization whenever an update is necessary.

The concept of media conversions [4] suits the next-generation mobile and wireless systems, as seen in the developments of third generation systems such as the European Universal Mobile Telecommunications System (UMTS) [2], [9] and the efforts of the Third Generation Project Partnership (3GPP) [3], [27]. For
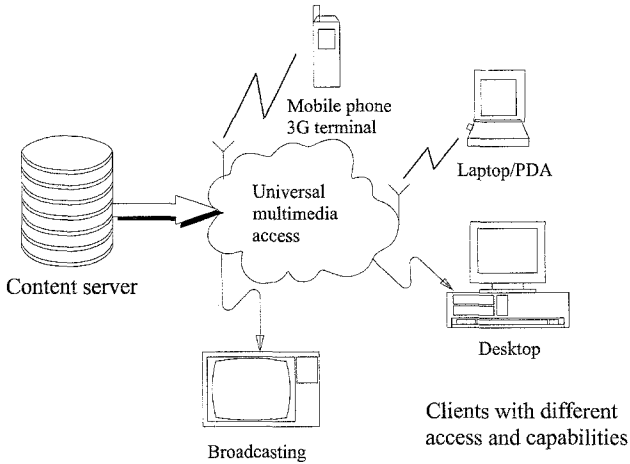
Figure 1. The UMA concept.

these applications, UMA will enable users access to future services independently on their choice of access technology, terminal equipment, and usage preferences.

2. The UMA concept

The UMA concept [5], [1], [7] is visualized in Figure 1.

UMA is truly multidisciplinary, involving multimedia signal processing, digital communications, radio systems, and telematics. Multimedia signal processing is essential in ensuring that the transmitted information is suitable and prepared for the communications system. Digital communications is essential in that the information has to be compressed to save bandwidth, high-level channel encoded for protection against noise in the channel, modulated to suit the physical channel, and multiplexed to allow for multiple access of the shared resources. This information passes through the radio system and is input to the network. Assuming a packet network, it will deliver the packets to the receiver according to the given protocol stack. The packet is affected by the network traffic and also by the quality of service (QoS) scheme defined for this specific network and client. Finally, the information is presented to the viewer at the receiving end. The UMA concept covers the universal access to multimedia information through this complete system.

The multimedia content for UMA-enabled systems first needs to be analyzed and then possibly converted to suit the different telecommunications requirements and access technologies. Analysis is the part of the system that describes the multimedia content, and the conversion is the part that adapts the content to different client capabilities given by the terminal and the telecommunication sys-
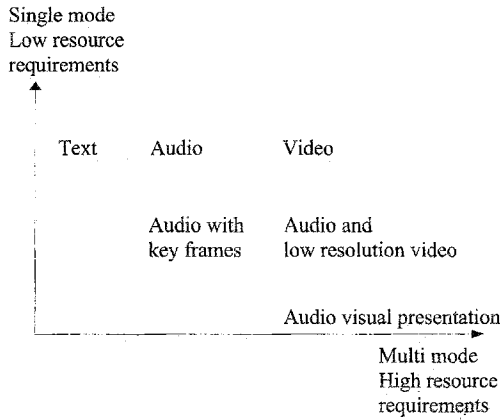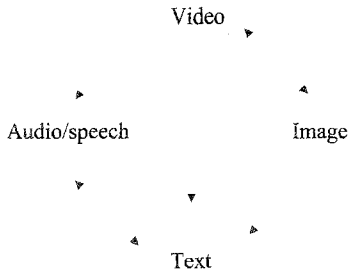
Figure 2. UMA media conversion space.

Figure 3. Modality changes for an audiovisual signal.

tem. Content conversions can be done prior to a client request or instantaneously subsequent to a request, depending on network configurations and protocol use.

3. Content servers for UMA-enabled systems

Figure 2 shows several different types of some of the possible media conversions you would expect a UMA client to be able to receive.

To be able to adapt this content to the different terminals, we can take two approaches. The first is to create conversions of the visual information by scaling the spatial or temporal resolutions. The second alternative is more complex and includes converting from one multimedia format (or modality) to another. Figure 3 shows possible modality changes. A multimedia presentation usually consists of combinations of the four main modalities; video, audio/speech, text, and images.

Channel bandwidth requirements will vary for different modalities. A video sequence usually has higher channel requirements than a presentation containing

Table 1. Typical client requirements for different modalities

Modality	Typical	Requirements
Video	Original video including audio	2 240 kbps, time critical
Images and audio	20 images and CD-quality audio	240 kbps audio, time critical 500 kbps visual, non-time critical
Images and speech	10 images and telephone quality speech	8 kbps speech, time critical 250 kbps image, non-time critical
Images and text	Two A4 pages of news including imagery	100 kbps, non-time critical
Text only	Four A4 pages	10 kbps, non-time critical

only text. This can be made scalable by combinations of different modalities. All of these possible combinations will require some conversion of the original multimedia content along with techniques to change from one modality to another. These new multimedia presentations will, like the original presentation, be time critical. The audio/speech sequence requires a minimum bandwidth to avoid stops in the sound. Images and video require a minimum bandwidth and channel quality to ensure QoS and also to meet client preferences on time delays [22]. The delay relates to the size of the imagery, which must, in one way or another, match the channel capabilities so that the transmission time does not exceed a certain limit. This limit will specify how long the client is willing to wait to get all the images and will be a trade-off with the received quality of the content.

Examples of modality changes are bandwidth reduction by speech extraction from the audio sequence and presenting this as a text string, extraction of a number of key frames of a video sequence, adapting the spatial size of the content, varying compression ratios and/or color schemes, or audio quality adjustments. Conversion from video to images and text includes almost the same operations as conversion from video to images and audio, as well as a speech-to-text conversion or a text-summarization process. Conversion tools range from key frame extraction, speech recognition, transcoding, and speech and image recognition summary generation based on the information given in the media descriptions.

Table 1 gives some examples of typical channel requirements for different modalities.

3.1. Viewer requirements

The UMA concept will place certain requirements on the multimedia viewers, such as:

(a) Being able to buffer a given amount of data to prevent frame delays during small channel variations when the channel characteristic changes dynamically. Commercial viewers use this today; the problem is how to take control of the buffer based on dynamic channel and network feedback.
(b) Being able to use a media description annotation to automatically extract media conversions from an original sequence. This should be done instantaneously and continuously, or alternatively the media descriptor can be as simple as a pointer to the correct conversion of the content on the server.
(c) Being able to provide adequate control over the streaming of the content, including fast response to changes in channel bandwidth, automated presentation of changing frame rates, and zoom functionalities.
(d) Being able to provide adequate support for negotiation procedures during time variations in channel conditions and access schemes (also for initial setup).

3.2. Media conversions

One of the important issues in access to multimedia content across platforms and terminal capabilities is media conversion. It consists of converting the original data to different modalities and resolution versions. One can classify media conversion methods according to different aspects. For example, depending on the output of the conversion, the following methods could be identified.

(a) *Scaling.* Here the original data is transcoded, manipulated, and compressed so that its size and data rate are reduced. Representative examples are image size reduction, video frame-rate reduction [12], and color reduction.
(b) *Translation.* Here the different media objects (image, video, text) present in the content are transformed from one modality to another. Examples of translation include text-to-speech (TTS) conversion, speech-to-text conversion (speech recognition), video-to-image (video mosaicking [29]), and image-to-text (optical character recognition).
(c) *Summarization.* Here the most relevant information is extracted from the original data. Examples of summarization are key audio clips and key frame-based summary [1]. The key frames could be organized in a sequential or hierarchical way.

Another way to classify media conversion methods depends on where they are performed—at the client, at the server, or at the proxy. Because the computational power, link bandwidth, and the storage capabilities of the client devices are often

limited, the transformation of the content at the client side is usually impractical. The use of an intermediate proxy makes the conversion of the data more efficient. However, the content providers may not have control over how their content will appear at different clients. Finally, it is possible to perform the conversion at the server side. In this way, the content provider can control the converted media and avoid problems faced by the proxy approach such as copyright-related issues.

Furthermore, the media conversion can be performed offline or "on the fly". The proxy-based media conversions must be immediate. However, this requirement is difficult to meet due to high complexity, particularly conversions related to video. Mohan *et al.* [18] proposed to use a representation scheme called InfoPyramid, which describes the audiovisual data in a multimodal, multiresolution hierarchy. It is a server-based, offline media conversion method. The InfoPyramid delivers dynamically converted versions of the original data according to the user preferences and the client device capabilities. In [10] a proxy-based media conversion is presented, based on translations such as image compression and video frame-rate reduction.

4. Network and UMA deployment

Downloading multimedia data involves capability negotiation between clients and servers or clients and gateways. This type of negotiation is supported in many media standards. As an example, in a videoconference application, terminals exchange capabilities and agree on a common mode of operation. If H.323 [3], [4] is the protocol used in this application, the two terminals negotiate which format of the video coding algorithm (H.261 or H.263) and which of the spatial dimensions to use: 352×288 as defined for CIF or 176×144 as defined for QCIF. If one terminal supports both CIF and QCIF resolution, and the other supports only QCIF, naturally QCIF will be used. The audio codec chosen must also be negotiated, e.g., for G.729 or G.723.

If video is transmitted between groups of users with different terminal or network capabilities, the need for converting one type of compressed signal into another type of compressed signal format through transcoding arises. The device that performs this operation is called a transcoder. A transcoder can also adapt to user preferences and capabilities. For example, in UMA applications, some users might be able to receive low-frame-rate video, while others can receive high frame rates. The transcoder can convert the video accordingly. In other words, the high-quality video will still be exchanged between users with the appropriate capability.

An alternative to video transcoding is scalable video coding, which MPEG-2, MPEG-4, H.323, and H.263 support. Scalable video coding allows for recovery of appropriate subsets of video bit streams to generate complete pictures of resolution with the proportion of the bit stream decoded. For example, let us

assume that R1 is 64 kilobits/second (kbps) and R2 is 128 kbps. Then, the base layer in the encoded bit stream would be R1, and the enhancement layer, containing higher-resolution information, would be R2.

The problem with scalable video coding is that it adds complexity to the system because it must produce as many layers as different preferences and clients exist in the network. Usually, a limited number of layers are generated, which might cause problems in a UMA application where various clients can coexist. On the other hand, video transcoding adds complexity and delays in the network, which might create a problem in real-time communication systems. It is evident that a combination of scalable coding and transcoding must coexist for efficiency.

5. Protocol requirements and network architectures for UMA

5.1. UMA protocol requirements

The UMA concept deployed in a heterogeneous network environment leads to what can be considered as "smart content" and hence will lead to "smart content servers". This means that network operators and service providers will be able to tailor content from an original source to a plethora of user devices with possibly widely varying characteristics as well as different network access communication links. New or modified protocols are required to enable full exploitation of the envisaged smart content.

Existing protocol mechanisms are available [11], [13], [14], [25] wherein a mapping can be determined between applications, the user terminal devices, and network requirements. For instance, the ITU H.323 protocol [13] provides a foundation for multimedia communications across IP-based networks (such as the Internet). Indeed, H.323 is designed to run on top of common network architectures, and, more importantly, it is not tied to any hardware or operating system. H.323 encompasses four major elements of a networked-based communications system [13]: terminals, gateways, gatekeepers, and multipoint control units. Note that all H.323 terminals must also support H.245 [14]; the latter is used to negotiate channel usage and capabilities. Similarly, the Internet Engineering Task Force (IETF) has developed the Session Initiation Protocol (SIP) [11], [25], which is an application-layer control signaling protocol for creating, modifying, and terminating multimedia sessions with one or more participants.

The above mapping and negotiation is usually established during connection setup and remains in place throughout the life of a connection. SIP provides the capability to change the parameters of an existing session; however, such a change is targeted at the case when a third party wishes to join or exit a multicast session. Ultimately, the performance of these existing protocol mechanisms [11], [13], [14], [25] depends on stochastic dynamic variations of telecommunication

channel characteristics, such as errors, delay jitter, and network congestion. In general, this problem has been solved by conservatively dimensioning the bandwidth buffer requirements as well as by the deployment of traffic management mechanisms. An implicit assumption is that the content remains fixed, and the objective is always to deliver the original content at all times. The fallacy in this approach is that there is wide recognition whereby users will demand access to content irrespective of their terminal characteristics and communication infrastructure. In other words, there is a natural demand for content even if it is scaled down from the original version; for instance, even if some of the individual media streams associated with a particular multimedia content source are not maintained through the media conversions before final delivery. This recognition represents the impetus for the development of media descriptions and hence UMA; that is, the notion that valuable information can be derived from a variety of conversions of a multimedia content source.

The development of smart content and smart content servers alone will not solve the issue of delivering the required flexible service to customers. New or modified protocols are essential and necessary to first establish and then map the various requirements between the user terminal equipment, the communication network access bandwidth, and other limitations, and finally, the optimal level of media conversions if deemed necessary. In other words, content-level parameters need to be taken into account. Further, and perhaps more importantly, the protocol is required to monitor the channel conditions such that an adequate and acceptable level of service is maintained through the life of the session. The latter requirements raise a number of interesting issues, e.g.:

1. What is the most appropriate time scale for monitoring of channel parameters such as errors, delay, and delay jitter? The impact that these parameters will have on the various levels of content delivery is unknown. The necessary interactions between the various communication layers will also need to be determined. For instance, the application layers will need information about lower layers and channel conditions in order to properly and adequately cater for the content variation of the media conversion.
2. There is a need to establish the threshold levels whereby media conversion is deemed necessary. Critical to this is the issue of response time and inherent communication delay between the user and the server within the network. There will also need to be a negotiation between network requirements and specified usage preferences.
3. In terms of deployment, consideration should be given to the media conversion scalability and dimensioning of the UMA content server to the available infrastructure.

These issues largely relate to QoS and what this means from a UMA application point of view. The UMA concept attempts to provide some mechanisms for maintaining a certain level of QoS even as network and other user behavior
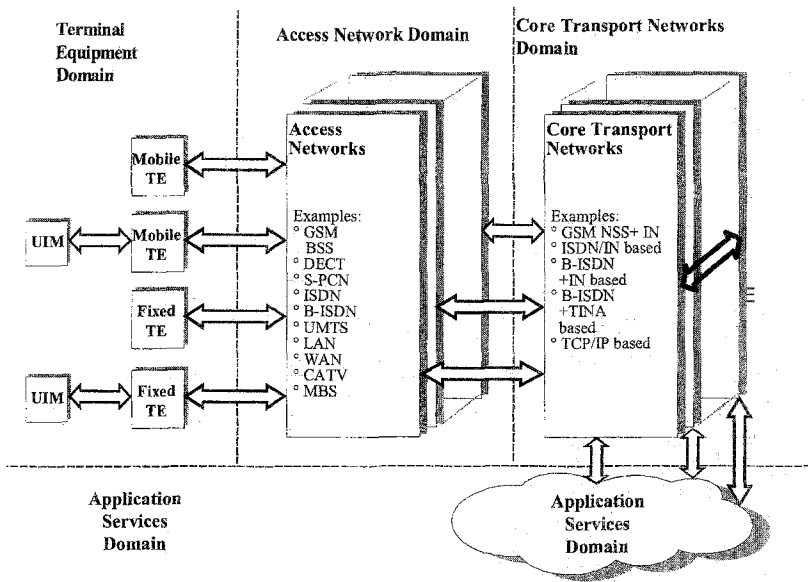
Figure 4. GMM communications system: Conceptual model of network architecture.

changes. However, tailoring QoS to individual multimedia content sources and usage requirements remains an open research question that is being addressed within the UMA test bed.

Finally, human interface factors such as user requirements and usage preference need to be addressed. The process of capturing a particular user's preferences and requirements is an important area for further consideration. For example, the knowledge of whether a user has a preference for a certain modality of the multimedia content, say audio rather than visual, has a strong bearing on the choice of media conversion. How this information is captured and whether it changes depending on the application clearly represents an open issue.

5.2. Network architectures for UMA

One capability of UMA is to provide a seamless interface between the wired and the wireless world. UMA makes this hidden from the client and extends the client's capabilities of staying connected regardless of access point and also provides the necessary mobility. This is very much along the lines of international work on the next-generation infrastructure, such as the Global Multimedia Mobility (GMM) work of ETSI [8] and also in the development of the Global Information Infrastructure (GII) [16].

The GMM architecture is shown in Figure 4 (from [8]).

This relates closely to UMA, as it will be supported, not by one system, but by a

set of systems and subsystems working together to provide the UMA capabilities. The GMM communications systems can be built from different combinations of current and future systems and subsystems. Therefore, the approach is not to define a single unique architecture but rather a framework within which different elements can coexist and work together.

It is stated in [8] that the network architectures for GMM communication systems should meet the following objectives:

(a) Cope with a diversity of services to be supported.
(b) Enable service provision by a large diversity of network operators and service providers ranging from global carriers to local specialized service providers.
(c) Allow the same application to be transparently accessible from different terminal equipment.
(d) Offer services in different environments, depending on local conditions.

Therefore the GMM network architectures comprise four conceptual domains (Figure 4). These are:

(1) Terminal equipment
(2) Access networks
(3) Core transport networks (including intelligence)
(4) Application services

UMA communication systems can be built in various ways using different combinations of elements from the four domains introduced above. As an example, consider the scenario where the user is involved in a training-on-demand program at work and has his desktop PC at work and a mobile device with a small screen and 144 kbps access available while traveling home. The program demands that work be continued outside the office and thus the GMM structure and UMA concept will ensure that the interface between the terminal equipment domain and access network domain in Figure 4 is seamless by the use of media conversions based on the description of the content. The interface between the access network domain and the Core transport networks domain is dependent on the available infrastructure and will be an input parameter in the description of the content, e.g., by giving hints on the transcoding to be used depending on error or delay characteristics. The final interface to the applications services domain will reflect on the choice and design of the multimedia viewers.

Issues to be determined are on the use of the Internet Protocol (IP) environment in the architecture, the envisaged user requirements, and influence of their usage preferences (human interface factors), scalability and transcoding for network deployment (media conversions), adequacy of content media bases and the placement of these.
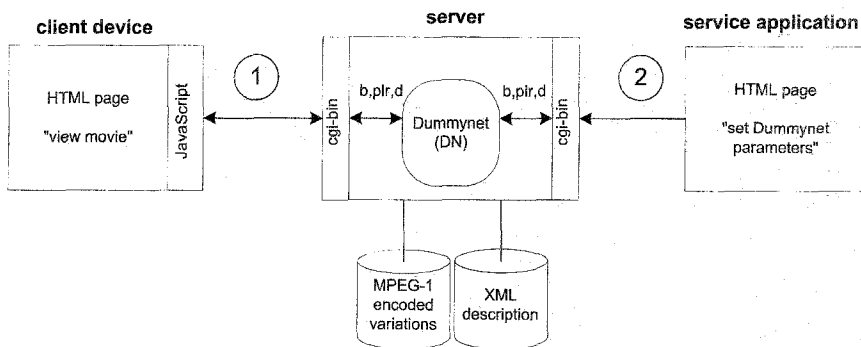
Figure 5. Test bed for validating the UMA concept.

6. UMA test bed

A UMA test bed is under development [2], [17], [23], [28], intended to serve as a media research engine for the UMA concept. It consists of a server including several conversions of MPEG-1-encoded content, a network simulation application, and three client computers. The different media are described using the MPEG-7 description schemes coded in XML [19]–[21].

The client/server network environment for the UMA concept is visualized in Figure 5.

The router in the network is running DummyNet (DN) [6]. DummyNet is a flexible tool for bandwidth management and for testing networking protocols such as IP. DummyNet works by intercepting packets in their way through the protocol stack and passing them through one or more channels that simulate the effects of bandwidth limitations, propagation delays, bounded-size queues, packet losses, etc. Each channel can be configured separately, and packets are forwarded to the appropriate channel using a packet filter. Hence we will be able to apply different limitations/delays to different traffic according to the specified filter rules (dynamic variations can be applied to channel bandwidth, propagation delays, and packet loss; these can be either deterministic or statistical).

6.1. Data access

The content server is currently running an Apache Hypertext Transfer Protocol (HTTP) server allowing the content to be easily accessed by the clients. The test bed operation consists of two Hypertext Markup Language (HTML) pages, which connect to the server/DummyNet through a JavaScript and plain HTML code, respectively. It also includes two different scripts, which set and read the DummyNet parameters.

The media conversions are accessed from the clients through an HTML page.

The most suitable content for the client's capabilities is selected through a JavaScript in this HTML page. This page performs a search through the content descriptions and presents the correct one to the client. Control of the DummyNet parameters is with an HTML page, which sets the transport condition through a script.

The dynamics of the test bed are through signal flows 1 and 2, as indicated in Figure 5. Signal flow 1 is the signal flow between the service application and the server, whereas signal flow 2 is the control signal flow setting the DummyNet parameters (bandwidth, packet loss ratio, and delay). After these parameters are set, and whenever they change, they are sent to the server through a script. The client device requests for the bandwidth parameter prior to presenting the content.

When the user requests for content, a script is invoked, which requests the actual bandwidth parameters. The parameters are read from the server, and a new HTML page is created instantaneoulsy, where a JavaScript requests for the XML code containing the media descriptions. The search decides which media conversion is the most suitable for the client. For example, this could be the conversion with the highest fidelity measure, but still meeting the channel constraints as read from DummyNet. The JavaScript finally sends a request to the server for the selected content, which is played back to the user through the Windows Media Player (as an ActiveX component in the HTML page).

6.2. Viewers

Each of the UMA clients can run a viewer. Currently, in the test bed, it is possible to stream MPEG data, i.e., play the movie while it is downloading, rather than first download the whole movie and then play it. The clients are currently running the InterVU player [15] and the Windows Media Player [30] for MPEG-1 encoded data.

6.3. Example of MPEG-7 annotations using the Variation Description Scheme

Figure 6 shows the use of the proposed MPEG-7 Variation Description Scheme within our test bed [23].

The concept of Figure 6 assumes a mechanism within UMA for negotiation of client terminal capabilities—in this case, its available bandwidth. This provides the parameter vector b_i, which is input to the description schemes. The bandwidth b_i is then compared to all of the available variations of the content through comparing b_i to the description for media conversion P_m. The media conversion will access the original MPEG-1 encoded content and output P_m, which is the most suitable conversion of the original content according to the client's terminal bandwidth.
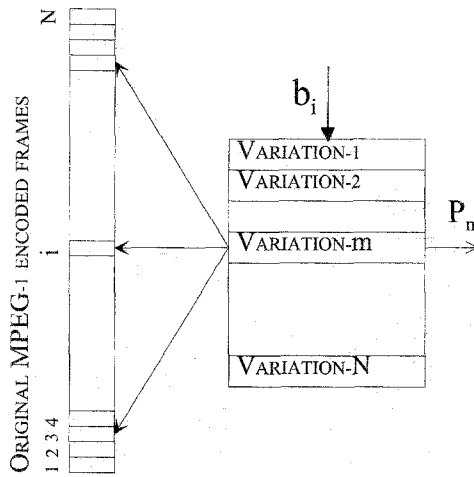
Figure 6. Use of the Variation Description Scheme in the UMA concept.

In the example, the media conversion is in the form of an extraction of key frames according to their significance in the sequence [3], [4]. The UMA test bed is set up to dynamically change the bandwidths to each of the clients connected to the network to be able to simulate the expected dynamics of a real scenario. The media conversion will at any given instance pick these frames from the original MPEG-1 encoded content database for transmission to the client.

7. Conclusions

UMA gives the basis for the future of efficient multimedia communications, providing the concept of "Multimedia anytime, anywhere". UMA provides a seamless access to multimedia information over wired and wireless systems. The merger of telecommunication systems with the MPEG-7 Descriptors and Description Schemes of multimedia content provides the mechanisms for maintaining an effective connection regardless of communication resources. MPEG-7 enables storing the content as a single entity, using the descriptions to select the appropriate content for the client's capabilities. The mechanism is based on media conversions, selecting a conversion suitable for the given requirements.

A test bed is under development, providing insight on the following issues for future research:

- Simulations with visual quality comparisons between different conversions for different client terminal capabilities, necessitating tables with descriptions of content according to certain channel conditions using conversions

of the original content. These descriptions have to be formalized, e.g., by using the DDL as proposed by MPEG-7.

- Necessary requirements for multimedia viewers in order to handle the concept of UMA.
- Requirements on content servers, noting that content no longer has to be stored on servers in a large set of formats suitable for any client. In fact, media conversion operations will be performed in transcoders that reside in gateways in the network. The main idea is that an MPEG-7 client will exchange its capabilities using the MPEG-7 descriptions, thus negotiating conditions with the server on establishment of connection and revising these during the connection.
- Necessary UMA protocol requirements establishing a mechanism for ensuring seamlessness when moving from one environment to the next, such as going from a wired access to a wireless access while viewing multimedia content.
- Discussions on a UMA architecture linking this to the GMM architecture and GII. This architecture allows for an evolution of existing systems in coexistence with new systems. The new systems just act as new additions in the flora of building blocks available for the user at any time and location.

References

[1] Y. Abdeljaoued, T. Ebrahimi, C. Christopoulos, and I. M. Ivars, Video summarization for universal multimedia access applications, ISO/IEC/JTC1/SC29/WG11 M5105, 1999.

[2] A. B. Benitez (Columbia University), J. R. Smith (IBM), J. Chicharo (University of Wollongong), A. Perkis (NTNU), S. Sull (Korea University), C. Christopoulos (Ericsson), and T. Suzuki (Sony), Report on core experiment on media transcoding hint DS, ISO/IEC JTC1/SC29/WG11/ M5803, 2000.

[3] N. Bjørk and C. Christopoulos, Transcoder architectures for video coding, *IEEE Trans. Consumer Electronics*, 44(1), 88–98, February 1998.

[4] N. Bjørk and C. Christopoulos, Transcoder architectures for video coding, in *Proceedings of IEEE International Conference on Acoustic Speech and Signal Processing (ICASSP 98)*, vol. 5, Seattle, WA, pp. 2813–2816, May 12–15, 1998.

[5] C. Christopoulos, T. Ebrahimi, V. Vinod, J. R. Smith, and R. Mohan, MPEG-7 application: universal access through content repurposing and media conversion, ISO/IECJTC1/SC29/WG11 M4433, 1999.

[6] DummyNet, www.iet.unipi.it/luigi/ipdummynet/, implemented in FreeBSD from 2.2.8 RELEASE and 3.1 RELEASE CDs.

[7] T. Ebrahimi and C. Christopoulos, Can MPEG-7 be used beyond database applications, ISO/IEC/JTC1/SC29/WG11/M3861, MPEG99, 1998.

[8] ETSI report, GMM—Global Multimedia Mobility, a standardisation framework for multimedia mobility in the information society. June 13, 1996, Programme Advisory Committee Approved March 1, 2001.

[9] European Telecommunications Standards Institute, www.etsi.org Technical bodies: UMTS and/or SMG.

[10] A. Fox, S. D. Gribble, Y. Chawathe, and E. A. Brewer, Adapting to network and client variation using active proxies: Lessons and perspectives, *IEEE Personal Comm.*, 40, 1998.

[11] M. Handley, H. Schulzrinne, E. Schooler, and J. Rosenberg, SIP: Session Initiation Protocol, Internet draft, RFC2543, IETF, March 1999.

[12] J.-N. Hwang, T.-D. Wu, and C.-W. Lin, Dynamic frame-skipping in video transcoding, in *Proc. IEEE 2nd Workshop on Multimedia Signal Processing*, pp. 616–621, 1998.

[13] International Telecommunication Union, Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service, Recommendation H.323, Telecommunication Standardisation Sector of ITU, Geneva, Switzerland, January 1998.

[14] International Telecommunication Union, Control protocol for multimedia communication, Recommendation H.245, Telecommunication Standardisation Sector of ITU, Geneva, Switzerland, February 1998.

[15] InterVU MPEG-1 Player, Plug-in for Netscape, available through INTERVU Multimedia Manager, www.intervu.net, November 17, 1999.

[16] ISO/IEC JTC 1/SWG-GII Special Working Group on Global Information Infrastructure Report, ISO/IEC JTC 1 GII Roadmap: guidelines for evolution, management and development of GII standards, available at http://www.globalcollaboration.org/.

[17] S. Jensen, Evaluation of universal multimedia access using MPEG-7 description schemes, Masters thesis, Norwegian University of Science and Technology, Trondheim, Norway, February 2000.

[18] R. Mohan, J. R. Smith, and C.-S. Li, Adapting multimedia Internet content for universal access, *IEEE trans. Multimedia*, 1, 104–114, March 1999.

[19] MPEG-7 Requirements Document, V.10, ISO/IEC/JTC1/SC29/WG1 M2996, 1999.

[20] MPEG-7 Experimental Model (XM), ISO/IEC JTC1/SC29/WG11 M3112, 1999.

[21] MPEG-7 Description Schemes (WD version 1.0), ISO/IEC/JTC1/SC29/WG11 M3113, 1999.

[22] A. Perkis and D. G. Cardelo, Transmission of still images over noisy channels, in *Proc. of ISSPA'99*, pp. 789–792, Brisbane, Australia, August 23–25, 1999.

[23] A. Perkis (NTNU), J. Chicharo (University of Wollongong), S. Jensen (NTNU), C. Christopoulos (Ericsson), and Y. Abdeljaoued (EPFL), Report on validation experiments for universal multimedia access (UMA), ISO/IECJTC1/SC29/WG11/ M5364, 1999.

[24] Public documents for MPEG-7, drogo.cselt.stet.it.

[25] J. Rosenberg and H. Schulzrinne, Reliability for provisional responses in SIP, Internet draft, RFC2060, IETF, January 2000, work in progress.

[26] J. R. Smith, C.-S. Li, A. Puri, C. Christopoulos, A. Benitez, P. Bocheck, S.-F. Chang, T. Ebrahimi, and V. Vinod, MPEG-7 content description for universal multimedia access, ISO/IEC JTC1/SC29/WG11/M4749, 1999.

[27] Third Generation Partnership project, www.3gpp.org.

[28] Validation experiments for universal multimedia access (UMA) - ISO/IECJTC1/SC29/WG11 N2971, 1999.

[29] H. Wallin, C. Christopoulos, A. Smolic, Y. Abdeljaoued, and T. Ebrahimi, Robust mosaic construction algorithm, ISO/IEC JTC1/SC29/WG11 M5698, 2000.

[30] Windows Media Player, version 6.4.05.0809, www.microsoft.com/Windows/ mediaplayer/, November 17, 1999.