MPEG-4 TRANSMISSION OVER WIRELESS NETWORKS

Ralf Schäfer, Christoph Fehn and Aljoscha Smolic
Heinrich-Hertz-Institut für Nachrichtentechnik Berlin GmbH
Einsteinufer 37, D-10587 Berlin, GERMANY
Tel: +49 30 31002-560, fax: +49 30 3927200
e-mail: schaefer@hhi.de, fehn@hhi.de, smolic@hhi.de

ABSTRACT

This paper reviews some application scenarios for transmitting MPEG-4 encoded video over wireless networks. These include side information or interactive services in digital terrestrial video broadcast (DVB-T), audio-visual information over in-home networks and digital video over DAB (digital audio broadcast system). For the first two applications the "Delivery Multimedia Integration Framework (DMIF)" is investigated, while for the latter application encapsulation of MPEG-4 streams in MPEG-2 transport streams (TS) and appropriate pre-processing are considered.

1 INTRODUCTION

MPEG-4 is supposed to become a world wide standard for future interactive multimedia services and applications. Among those, services and applications over wireless networks will play a more and more important role. Two major classes of applications , i.e. „broadcast" and „real time communications" and three major classes of wireless networks, i.e. „digital broadcast networks (DAB[1], DVB-T[2])", „communication networks (DECT, UMTS)" and „digital in-home networks (IHDN[3])" will play a dominant role in this respect. The work at HHI mainly concentrates on broadcast and communication applications over broadcast and in-home networks.

As far as broadcast networks are concerned, it is clear, that MPEG-2 compression will be used for digital terrestrial television according to the DVB-T standard. However, MPEG-4 encoded material may be embedded in the private data streams as interactive add-on to the regular programme. Within DVB network independent protocols have been defined in order to carry out IP based interactive applications over the network. A possible protocol stack for this scenario consists in embedding MPEG-4 into RTP/UDP/IP and to embed the IP stack via the DSM-CC data caroussel into the private section of MPEG-2 Transport Streams. Another possibility is to use the "Delivery Multimedia Integration Framework (DMIF)", which is currently under development in MPEG, for this application.

For the transmission over wireless in-home networks, also MPEG-2 transport streams will be used. For communication purposes (e.g. videophone on PC within the home) or for retrieving MPEG-4 data from a local home server, DMIF is also a suitable technology. Therefore DMIF is currently implemented at HHI in order to evaluate its performance in some real time and retrieval applications over the above-described wireless networks.

For video broadcast over DAB only small data rates of about 1 Mbit/s are available. This is not sufficient for the transmission of 625 (576 visible) lines, 50 Hz TV material. Therefore investigations have been carried out to down convert the TV material, to encode the video using MPEG-4 and to up convert the decoded signal in the receiver to the display format[4]. For transmission, the MPEG-4 FlexMux streams are embedded into MPEG-2 transport streams as will be defined by MPEG.

This paper is organized as follows. In Section 2 the Delivery Multimedia Framework will be described in more detail and in Section 3 results on the pre-processing for MPEG-4 transmission over DAB will be presented.

2 THE DELIVERY MULTIMEDIA FRAMEWORK OF MPEG-4

DMIF (Delivery Multimedia Integration Framework) allows applications to transparently access and view multimedia streams whether the source of the stream is located on an interactive remote end system, the stream is available on broadcast media or it is on stored media. By adopting QoS (Quality of Service) metrics which relate to the media and not to the transport mechanism, DMIF hides delivery technology details to applications. These unique features of DMIF give multimedia application developers what they need in terms of permanence and richness beyond what is possible with each individual delivery technology.

In order to fully reach this goal, DMIF provides the DAI (DMIF Application Interface), a semantic API that allows the development of applications transparent to the supported delivery technologies. By using the DAI application developers can begin to invest in commercial multimedia applications with the assurance that their investment will not be made obsolete by new delivery technologies.

Figure 1 shows the DMIF communication architecture for the three main MPEG-4 services, namely broadcast, local storage and remote interactive scenarios.

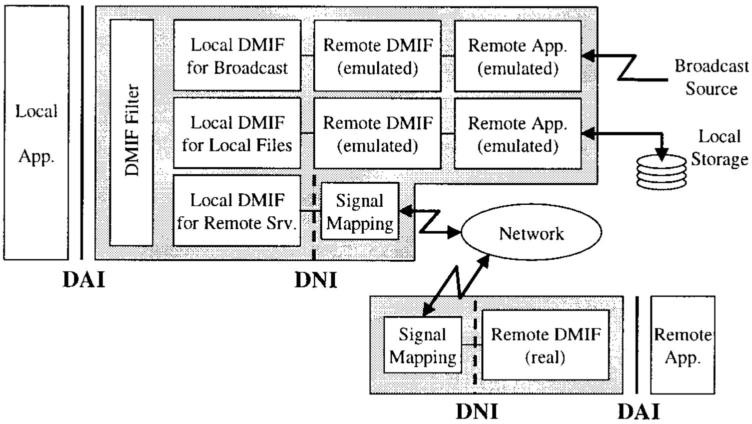

Figure 1: DMIF communication architecture

In the case of remote interactive services interoperability between end-systems is achieved by the definition of a semantic DNI (DMIF Network Interface), which abstracts the signalling between DMIF peers irrespectively of the supported delivery technologies. The parameters conveyed through the DNI are then normatively mapped onto network dependend native signalling (e.g. H.245 and H.323) when possible. Parameters which can not be mapped onto native signalling are carried opaque to the native signalling in which case a default syntax defined by MPEG will be used.

DMIF allows the transparent transmission of multimedia data using a variety of existing transport protocol stacks such as RTP/UDP/IP, MPEG-2 TS, ATM or H.223. For protocol stacks with fixed length packets (e.g. MPEG-2 Transport Stream) or with high multiplexing overhead (e.g. RTP/UDP/IP) an optional, low complexity multiplexing tool (FlexMux) is defined that allows interleaving of data with low overhead and low delay. This is particulary important for low bitrate applications.

As part of the implementation of an interactive client-server scenario, HHI is involved in the evaluation and standardization of DMIF, mainly through the verification of the DMIF interfaces. The client-server application uses the default signalling syntax for the communication between DMIF peers and allows access to MPEG-4 data over IP based networks, which include the DAB, DVB-T (via private data streams) and in-home networks mentioned above. A UDP/IP transport protocol stack is used for the transmission of the MPEG-4 encoded multimedia information.

3 MPEG-4 VIDEO OVER DAB CHANNELS

Figure 2 gives an overview of a possible system for MPEG-4 broadcast over DAB. The incoming video is in ITU-R. 601 format with a raw data rate of 170 Mbit/s. These data shall be coded using MPEG-4 and transmitted over a DAB channel with 1 Mbit/s. Besides the direct coding of the video signal it is possible to reduce the number of samples to be coded by down-conversion into an other format with a lower temporal and/or spatial resolution. At the receiver the decoded video has to be adapted to the display. Here it is planned to support a variety of future mobile multimedia terminals as TV-type receivers or laptops. This flexibility requires a video format conversion in the receiver.

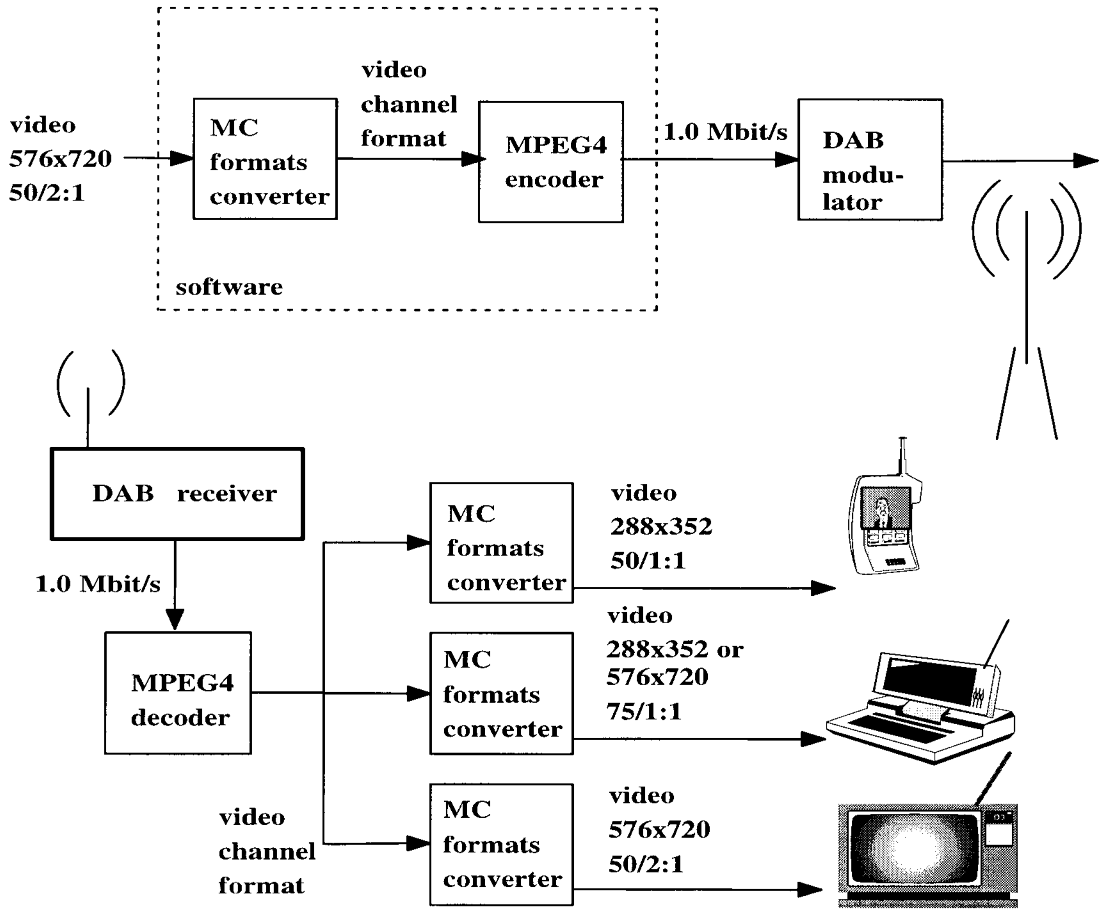

Figure 2: Overview of the MPEG-4 / DAB system

The process of down-conversion at the transmitter and up-conversion at the receiver produces conversion artifacts which increase with the decrease of the temporal and spatial resolution of the channel format. On the other hand the coding artifacts increase with the temporal and spatial resolution of the channel format. This is a trade-off, which is subject to optimization that has to include the whole system.

The potential channel formats that were evaluated are listed in table 1. The interlaced sequences were generated simply by low-pass filtering and sub-sampling of the ITU sequences. The progressive channel formats were generated using a motion compensated de-interlacing (MCD) technique followed by low-pass filtering and sub-sampling (and frame skip in case of CIF25).

Name	Frame size	Frequency [Hz]	Scan
ITU	704x576	50	Interlaced
¾ ITU	528x576	50	Interlaced
½ ITU	352x576	50	Interlaced
¼ ITU	352x288	50	Interlaced
CIF50	352x288	50	Progressive
CIF25	352x288	25	Progressive

Table 1: Evaluated channel formats

The up-conversion is dependent on the channel and the display format. For this purpose a motion compensated interpolation (MCI) algorithm was developed. It consists of a hybrid recursive motion estimation, combining the advantages of a pixel-recursive (low complexity realization of the optical flow principle) and a block-recursive approach. A start vector is determined out of three candidates by a recursive block matching procedure taking into account a temporal and two spatial predecessors. The optimum candidate is the one with the smallest displaced block difference. Then the start vector is refined using a pixel recursive gradient method. The resulting dense motion vector field is used to interpolate unknown pixel values with a median filter that takes into account spatial and motion compensated temporal neighbors.

Simulations have been carried out including the whole system with down-conversion, MPEG-4 encoding and decoding and up-conversion to the different display formats. The results were judged in terms of subjective image quality by expert viewers.

Coding in ¼ ITU format resulted in a significantly lower image quality compared to all other formats. Too much information was destroyed through the down-conversion by a factor of two in both dimensions. The other extreme was the direct encoding in ITU R. 601 format. It was not possible to encode very complex sequences with a high amount of motion and much spatial detail in this format. Due to encoder overload several frames were skipped resulting in very annoying motion artifacts. The same holds for ¾ ITU, although less frames were skipped.

The best results for a display in ITU and in CIF50 were obtained using ½ ITU as channel format, followed by CIF25. The main reason is that with this channel format original image information is available at every time instant of the output sequence and no interpolation of intermediate images is necessary as for CIF25 channel format. The visual quality of intermediate images is always worse compared to spatially interpolated images. Although the decoded CIF25 sequence contains less coding artifacts, due to the lower raw data rate, the resulting visual quality of the converted ½ ITU sequence is better.

For a display in CIF75 the best results were obtained using CIF25 as channel format. Here ½ ITU has no advantage because no image information is available on two out of three time instances of the output sequence, i.e. the frames have to be generated fully from information from preceding and following fields. The same situation holds for CIF25 channel format, but this format has only half of the raw data rate compared to ½ ITU resulting in less coding artifacts.

4 CONCLUSIONS

We have presented some concepts for the transmission of MPEG-4 video over wireless networks. For real-time applications (e.g. videophone), for side information in DVB-T or DAB channels or for MPEG-4 retrieval over in-home networks DMIF is a powerful tool to access and view multimedia content. Concerning the concept of MPEG-4 video transmission to mobile receivers over DAB, several algorithms for video format conversion were briefly described. Finally, we presented experimental results on the optimization of the system by evaluation of several potential channel formats and format conversion algorithms.

5 REFERENCES

[1] W. Klingenberg, A. Neutel, "MEMO: A Hybrid DAB/GSM Communication System for Mobile Interactive Multimedia Services", Proceedings of ECMAST'98, Lecture Notes in Computer Science 1425, ISBN 3-540-64594-2, Springer-Verlag, May 1998, pp. 493-503

[2] R. Burow et al., "On the performance of the DVB-T System in Mobile Environments", Lecture Notes in Computer Science 1425, ISBN 3-540-64594-2, Springer-Verlag, May 1998, pp. 467-480

[3] P. Robertson, H.P. Huth, K. Fazel, O. Klank, W. Bauerschmidt, "A Sytem Design for a Wireless Home Multimedia LAN", Proceedings of ECMAST'98, Lecture Notes in Computer Science 1425, ISBN 3-540-64594-2, Springer-Verlag, May 1998, pp. 453- 464

[4] A. Smolic, K. Rümmler, J.-R. Ohm, R. Schäfer, S. Bauer, "MPEG-4 video transmission over DAB/DMB: joined optimization of encoding and format conversion", Submitted to MOMUC'98, Berlin, October 12-14, 1998.

6 ACKNOWLEDGEMENT

The work described in this paper is carried out within the framework of the MEDEA project A116 M4M (MPEG fo(u)r Mobiles) and within the MINT (Multimedia communiction on Integrated Networks and Terminals) project. The M4M project is a major European project concerned with MPEG-4 (hardware) implementation for real time application as well as for broadcast applications. In Germany the M4M project is funded by the national government. MINT is a German collaborative research project equally funded by the BMBF.