FACTA UNIVERSITATIS (NIŠ)
SERIES: ELECTRONICS AND ENERGETICS VOL. 7, NO. 1, 1994, 25-38

ADVANCED DIGITAL HDTV OVER ATM NETWORKS

Zoran Bojković, Slobodan Lazović and Miodrag Bakmaz

Abstract. This paper summarizes some recent effects and research activities in the field of the introduction of Advanced Digital High Definition Television AD-HDTV over Asynchronous Transfer Mode ATM-based networks. Some types of layered video coding schemes are identified. Taking into account that cell losses seem to be the major drawback for Variable Bit Rate VBR coding in ATM network, cell losses protection is analyzed and some solutions concerning this problem are presented. From the point of view of digital HDTV, hierarchical coding as well as Moving Pictures Experts Group MPEG compression are recommended

Key words: HDTV, ATM-based networks, MPEG, BISDN, packet transmission.

1. Introduction

Future telecommunication networks should be able to provide various information services as efficiently and economically as possible. These services will be provided through a single network where all network resources are shared and integrated standard access is provided. This network should provide communication capacity and maximum flexibility in the sense to be able to handle wide range of bit rates with different statistical natures. This is realized by dividing each information flow into short entities to which are attached short flow-identification lables and by transporting these entities to their destinations. This transportation mode is called the asynchronous transfer mode ATM. It is widely recognized as the solution for implementing

Manuscript received February 22, 1993.

The authors are with University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11000 Belgrade.

the broadband integrated services digital networks BISDN. ATM is essentially a packet oriented transfer mode using asynchronous time division multiplexing techniques. Namely the information is divided into the fixed-length transmission units called cells and is transmitted asynchronously.

Video signals are especially well-suited for packet transmission. Packet networks can carry variable bit-rate signals directly, no buffering or rate-control feedback is necessary at the transmitter. However, packet networks also provide some difficulties for video coders. When a packet is lost, usually more than a hundred bits are not available at the video decoder.

Because of large amount of data required to represent pictorial information, visual services are still very costly to implement. As a result, much time and money has recently been spent on the coding and compression of still and moving images in order to reduce the transmission costs in terms of required bandwidth and, hence, bring the widespread use of such services closer to reality. This, in turn, has led to the introduction of many new, high-compression video and image coding schemes over the last few years. The most successful of these schemes today have been the variable bit-rate VBR transform coding based algorithm which are able to provide efficient, high quality image transmission with compression ratios in excess of 20 : 1 by exploiting the flexibility of the ATM environment and allowing more or less information to be transmitted as the amount of activity in an image varies [1].

In this paper, we focus on Advanced Digital High Definition Television over ATM networks. In Section II ATM network characteristics are briefly described. Possibilities for laying the video data are addressed in Section III. The problem of cell loss, its effect on variable bit-rate coding technique as well as the corresponding protection are discussed, too. Digital HDTV approach is addressed in Section IV, while Section V summarizes the paper.

2. ATM for the Future roadband ISDN

The BISDN is a new concept for exchange area communications. Based on high-speed circuit and packet switching, BISDN will integrate services such as voice, data, video, Local Area Network LAN, Metropolitan Area Network MAN. Among various components of the BISDN, the cell-based ATM multiplexing technique is a powerful tool for providing integrated services effectively. The ATM technique has some desired attributes of both circuit switching and packet switching. Unlike the conventional packet switching ATM uses a fixed-length packet so that it is suitable for high-speed applications. Each fixed-length packet, called a cell, contains a header and an information field. The header allows easy multiplexing of user data and also, allows dynamic bandwidth allocation to accomodate users need. It typically includes the routing information in addition to some other fields, while the adaptation header in the information field carries the service and/or technique dependent information end-to-end. Since video services offer a large market for residential users, but at a significant higher bandwidth requirements than voice and data, success in adopting ATM for BISDN depends a great deal on the feasibility of transmission motion video in an efficient manner.

2.1. Main ATM features

The ATM is shown in Figure 1. Generally speaking, the information field is available to the user, while the header field carries information that belongs to the ATM layer functionality itself. Like packet switching techniques, ATM can provide a communication with a bit rate that is individually tailored to the actual need, including time variant bit rates.

Header	Information field
5 octets	48 octets

Fig. 1. ATM cell structure

In ATM based network, the multiplexing and switching of cells are independent of the actual application. The same modul of equipment in principle can handle a low bit rate connection as well as a high bit rate connection. On the other hand, the flexibility of the ATM-based BISDN network access due to the cell transport concept supports the idea of a unique interface which can be employed by a variety of customers with quite different service needs.

Major characteristics of an ATM-based BISDN include: high flexibility of network access, dynamic bandwidth allocation on demand, flexible bearer capacity allocation and independance of the means of transmission at the physical layer, easy multimedia integration (data/voice/image/video).

From a user's point of view, the advantages due to the introduction of ATM include:

- format independent, rate-free, and time transparent transmission,
- reduction of end-to-end delay,
- constant video quality even for active motion area/frames,
- quality control by the user instead of rate control by the channel

2.2 Video evolution strategies

One of the most important concept arrived to video communication is to adaption of the ATM as the switching protocol which will be used in the future BISDN. This network is the result of a growing demand on broadband services supported by the availability of high transmission speeds and switching as well as signal processing techniques.

HDTV will have a significant potential impact on future visual communication services and is expected to be one of the many different services provided by the communication network in the future. Such a network will represent a culmination of the current evolution in which different services will be offered with integrated access, transport, switching and network management.

2.3. Video service aspects

ATM provides the cell transfer for all services [2]. From the point of view of bit rate, the ATM-based network can be split into constant bit rate CBR and variable bit-rate VBR services. CBR maintains high transmission quality and can be simply controlled. On the other hand potential advantages VBR networks over conventional CBR digital switching channels include bandwidth saving, constant picture quality elimination of the smoothing buffer, flexibility to a wide range of different services and easy network planning.

In the ATM network, VBR transmission of video signals becomes possible by dividing video signals into cell and multiplexing those cells statistically. Therefore video signals with burstiness can be transferred maintaining high quality, even if the amount of encoded video information become large because of violent movement in the video scene.

BISDN will integrate both interactive and distribution services. The first services will consist for instance, of person to person communication (i.e. videophone), person or group to group communication (i.e. videoconference) and video surveillance at bit rates below 2 Mbit/s. The distribution services include digital TV, according to CCIR Recommendation 601, sometimes refered to as Extended Quality TV-EQTV and HDTV distribution at bit rates 45 and 135 Mbit/s, respectively.

2.4. Standardization efforts

In 1988, CCITT selected ATM as the transfer mode first set of Recommendations was agreed upon worldwide. This set of Recommendations specifies the details of the ATM basis, such as cell characteristics. The specifications complete the lower layers of the OSI Reference Model. The details of the higher layers, i.e. broadband signalling, are currently under study in CCITT and standards are expected in the 1993-1994 period [3].

Three standards have emerged which are facilitating the growth of new image communications applications. They are: the Joint Photographic Experts Group (JPEG) standard for still picture compression, the Consultative Committe on International Telephony and Telegraphy (CCITT) Recommendation H. 261 for video conferencing and the Moving Pictures Experts Group (MPEG) for fullmotion compression on digital storage media and digital communication. A common feature of these compression methods is the use of a combination of predictive and Discrete Cosine Transform (DCT) coding. They are used for temporal/spatial correlation reduction, with Huffman variable length codes VLC for statistical data reduction-entropy coding.

3. Layered coding for video data transmission

The layered coding is a coding procedure where the source information is divided into two or more layers following a criterion of spatial or temporal resolution, quantization or even combinations of the previous criteria. Research in packet video coding so far suggests that layered coding strategies are most suited for ATM. As a result of the current activities: in MPEG, a generic layered codec presented in Figure 2 will likely be the standard for all video communication applications [4]. One of the requirements in the standard is the so-called "scalability", which means the ability to decode images of different resolution and/or quality.

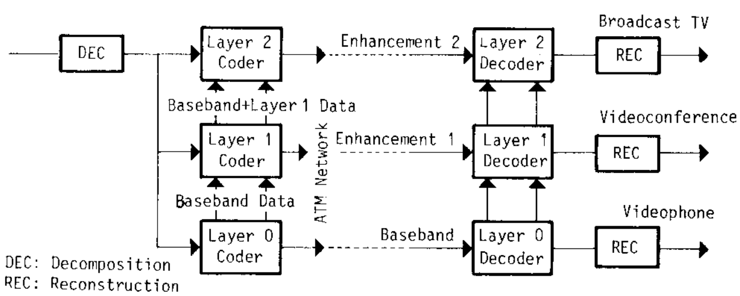

Fig. 2. MPEG layered codec.

One way of achieving scalability is by splitting the source image into different layers and encoding them separately. Generally, the layers consist of a base layer and other higher or enhancement layers. Besides scalability, the advantage of layered coding is the cell loss resilience in an ATM environment. As an example, consider a bloc of two-layer encoder and decoder shown in Figure 3.

The coded interframe video data are divided in two parts. The first part is the essential information about the picture and is transmitted through a guaranteed channel of an ATM network. The second channel represent the shared channel of the network. Cells from this channel can be lost. Increasing the portion of the bit rate to be transmitted in the guaranteed channel makes the codes more immune to the effect of lost cells from the second channel. The two-layer coder has to be designed in such a way that the base layer coder can generate pictures of acceptable quality, but with a minimum bit rate.

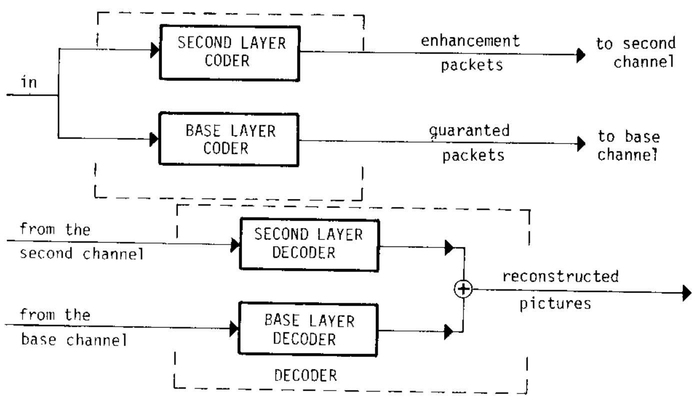

Fig. 3. Two-layer codec.

Layered coding techniques for video data transmission proposed so far can be classified into four basic categories, i.e.

- bit plane separation,
- frequency domain separation,
- combined bit-plane-frequency separation, and
- feature-plane separation

3.1. Bit-plane separation

Bit-plane separation represents the simplest way to layer an image since the video signal is digitized into a PCM format. A video frame quantized into B bits is separated into B bit-planes and most significant bit plane is assigned with the highest priority. Picture quality decreases with cell loss rate for bit plane separation.

3.2. Frequency-domain separation

In frequency-domain separatin, the pyramid data video structure is formed in terms of their frequency components. The picture quality generally does not degrade linearly with the cell loss rate. An example of a frequency-domain separation scheme is subband coding SBC. The basic idea of the SBC is to divide the frequency band of signal into a number of subband by a bank of bandpass filters. Each subband is then translated to baseband by down-sampling and encoded separately. At the receiver, the subband signals are decoded and up-sampled back to the original frequency band by interpolation. The signals are then summed to give a close replica of the original signal. The advantages of a subband coding scheme are: a) each band can select the optimum coding algorithm, b) quantization noise generated in a particular band is not allowed to spread to other bands and c) parallel processing can be applied to each band. Subband coding is very suitable for layering. In order to improve the coding efficiency, subband coding can be combined with other coding techniques like differential pulse code modulation DPCM, discrete cosine transform DCT, vector quantization VQ.

In motion compensated interframe SBC approach, the spectrum of each frame of video signal is first decomposed into smaller frequency bands where each can be coded. To preserve its hierarchical structure, each band is coded independed of higher frequency bands but can share information with the lower bands. Due to the fact that the nature of each decomposed band differs from one to another, it can be expected that the horizontal, vertical and diagonal bands generate different displacement estimates.

3.3. Combined bit-plane-frequency separation

In combined bit-plane-frequency separation, the advantages of bit plane separation in the high cell-loss rate and frequency domain separation in the low cell-loss rate are combined. The improvement can be seen if we note that the curve picture quality degradation versus cell loss rate moves closed to Shannon's distortion-rate function which defines the best quality subject to certain cell loss rates. An example of combined bit-plane-frequency separation is the combined-transform coding. This technique can be extended to variable bit rate interframe applications.

An interframe motion-compensated combined-transform coding with vari bit rate is shown as a block scheme in Figure 4.

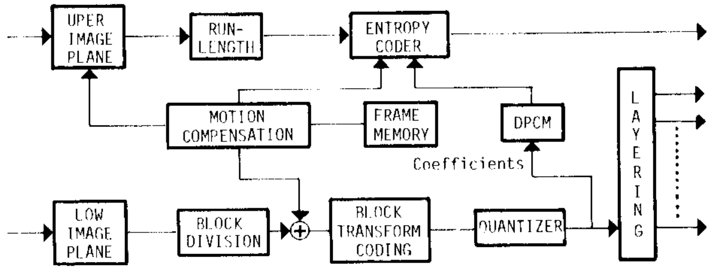

Fig. 4. Interframe motion-compensated combined-transform. coding with variable bit rate

A video frame is divided into the upper image plane and the lower image plane. The upper image plane contains the most important information which is run-length and entropy coded, together with the motion coefficients of the low image plane. The most important information is assigned as the highest priority. The lower image plane contains less important information which is block transform coded and assigned as the lower priority. The transform coefficients form a pyramid in the frequency plane. In the case of congestion, the cells in the bottom of the pyramid are dropped and performance degrades with the increase of cell loss.

3.4. Feature plane separation

Feature plane separation is a subjective separation scheme depending on the information interest for specific applications. Namely, in feature plane separation, the pyramid is organized in terms of the importance of features. For example, in one environment, the information of an object or about a part of considered object might be more important than the background.

3.5. Cell losses protection

When video signals are transmitted in packet mode through an ATM network, there exist several degradation factors like video signal time delay, bit errors in the transmission link, cell jitter and cell losses. There are three factors contributing to cell loss i.e. transmission bit errors, cell buffer overflows and excessive time delay. Cell losses inherent to the ATM network put specific requirements on the coding of services whether the rate is fixed or variable. It can be expected that the effect of cell loss will be very annoying and the resulting HDTV quality will not be acceptable unless some new protection scheme are implemented.

Many research activities have been conducted for cell loss protection. These include automatic repeat request, forced command refreshment, demand refreshment, selective recovery in the ATM adaption layer. The idea of layer source coding was also proposed. Different schemes including layered subband coding, layered transform coding and layered predictive DPCM coding, have been developed to minimize the performance degradation due to cell loss. The first step in the layered video service coding is to classify picture information in terms of its importance and then assign the most important class the highest priority. After the separation, the information is layered into a pyramid form and compressed by VBR coding schemes. In the decoder, these layered signals are decoded and combined to reconstruct the original signal subject to a certain distortion.

An experimental study of ATM video transmission for HDTV signals was described in [5]. The codec has an ATM cell loss compensation scheme using two-layered coding, block interleaving and data transposition with cell loss detection. Experimental results showed that the proposed cell loss compensation scheme improved the permissible error rate, while the high picture quality was kept in the ATM cell loss environment.

4. Digital HDTV Approach

High Definition Television HDTV will be the coming standard for television in the future. It will be one of the driving market forces for introducing fiber optic cables in homes, because construction of a broadband network is necessary to transmit wide-band HDTV signals. With the evolution toward a broadband digital network, HDTV transmission systems between broadcasting stations will be constructed in order to offer high-quality and reliable HDTV services to customers.

In HDTV transmission systems, image signals are expected to be transmitted betwen broadcasting stations or studios through codec links. Therefore, the capability for multiple encoding and decoding of signals without any visible degradation is necessary. Furthermore, the number of coding processes is limited because image signals should be processes at a very high frequency sampling rate.

4.1. Quality of HDTV and the human visual system HVS

The quality of the HDTV video we perceive is directly related to how the human visual system HVS responds to the video. The HVS can perceive detail more easily in the luminance than in the chrominance components.

Also errors in coding the high frequency components are not as visible as errors in coding the low frequency coefficients.

4.2. Coding of HDTV

The coding of HDTV for digital transmission via BISDN is of importance since digital fiber optics may represent the most viable means for avoiding transmission impairments and delivering very high image quality. For uncompressed video, a single bit error usually affects only one pixel which is not likely to be very noticeable. However for coded video data, a single bit error may affect a large picture area and this error may even propagate from frame to frame in the case of interframe coding. Furthermore, variable-length coding VLC is always applied in a digital compression system to exploit the statistical redundancy in the processed data. The use of VLC causes compressed HDTV to be even more susceptible to transmission errors since a single bit error may result in the loss of code synchronization.

4.3. Hierarchical coding

The BISDN will facilitate distribution of wide-band services like digital standard-resolution TV-sometimes referred to as Extended Quality TV and High Definition TV. As for HDTV, it will introduce wide-screen cinema quality at the television market by doubling the resolution of EQTV signal. These video services will most likely be transmitted at the H_2 and H_4 access rates of the BISDN of approximately 35 Mbit/s and 135 Mbit/s. Therefore, the bit rate of the EQTV signal has to be reduced by data compression to approxi mately 35 Mbit/s, while the HDTV bit rate must be reduced to approximately 135 Mbit/s.

Since EQTV and HDTV are different video services, data compression should be performed on the individual EQTV and HDTV signals. However, in the case that both services are covering an identical scene, the standard-resolution information is transmitted twice, once as the coded EQTV signal and once as part of the coded HDTV signal. Hierarchical or compatible coding schemes combine EQTV and HDTV coding in order to use the available network bandwidth as efficiently as possible. Principle of hierarhichical coding is presented in Figure 5.

The coding scheme consists of decomposition and reconstruction schemes as well as the coding and decoding sections. The decomposition section decomposes the HDTV complements signal HDTVc sometimes called enhancement signal. Given the EQTV signal, the latter can be used to reconstruct the HDTV signal. Each sub-signal is coded under supervision of

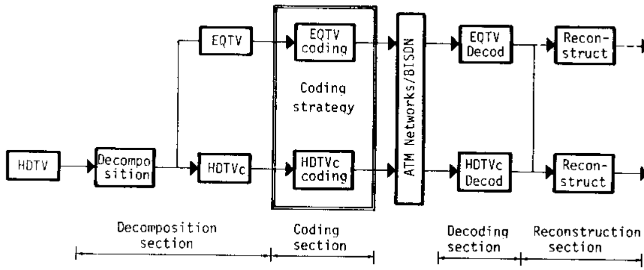

Fig. 5. Hierarhichical coding scheme

a certain coding strategy before transmission across the ATM-channel. The receiver may either receive, decode and reconstruct the EQTV sub-signal visualizing the signal at an EQTV monitor, or it may receive, decode and reconstruct both subsignals so that the signal can be displayed at full resolution HDTV monitor. Hierarchical coding shemes are well suited to ATM networks because the EQTV signal or packets can be transmitted with a higher priority than the HDTVc signal or packets. This mechanism will guarantee a reasonable low-pass reconstruction of the HDTV signal in case of network congestion and cell loss. The great advantage of such a hierarchical coding scheme is that it does not require any extra channel capacity for the low-resolution signals.

4.4. Key elements of Advanced Digital HDTV

Advanced Digital High Definition AD–HDTV is a layered digital system that consists of:

- MPEG + + video compression,
- MUSICAM audio compression,
- Prioritized Data Transport format, and
- Spectraly-Shaped Quadrature Amplitude Modulation

These key elements are combined in a layered digital approach that alows them to operate as an effective system. At the highest level, AD–HDTV has three primary layers at the compression, packet and transmission levels. Each primary layer consists of sublayers that generally corespond to the reference model.

4.5. MPEG compression

MPEG is an ISO standard for compressed video on digital storage media. MPEG compression is the basis for AD-HDTV.

MPEG video compression operations include motion-compensated predictive coding and adaptive DCT quantization. Motion compensated prediction assumes that "locally" the current picture can be modeled as a translation of the picture at some previous time. Locally means that the amplitude and the direction of the displacement need not be the same everywhere in the picture. The motion information is part of the necessary information to recover the picture and has to be coded appropriately [6].

Picture frames in a self-contained Group of Pictures GOP are classified into one of three types: intracoded frames (I-frames), predictively coded frames (P-frames) and bidirectionally coded frames (B-frames). GOP and their temporal prediction relationships are show in Figure 6.

I-frames use purely spatial compression and are processed independently of other frames. P-frames are coded by a motion-compensated predictive coder using the previons I or P-frames. B-frames are coded by a bidirectional motion-compensated predictive coder using the two adjacent I or P-frames. By the picture information in each frame we mean pixel values for I-frames and residual error after prediction for P-and B-frames. On the other hand, the picture information in each frame is organized into 8 $\times$ 8 blocks, transformed by DCT operation and then adaptively quantized. Each quantized DCT coefficient is assigned a fixed-length internal code and passed on to the priority encoder for variable length coding in the form of MPEG codewords. Each frame may be divided into slices, each slice consists of several macroblocks.

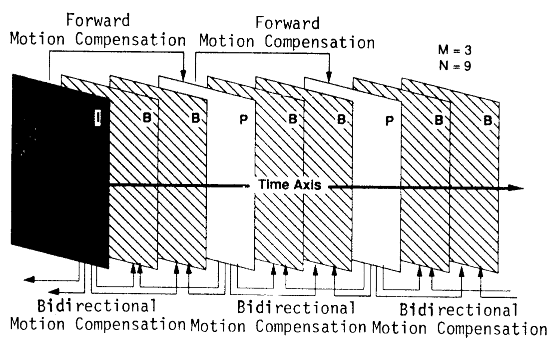

Fig. 6. Group of pictures classifications

The structure of MPEG implies that if an error occurs within I frame data, it will propagate through all frames in the GOP. Similarly, an error in a P-frame will affect the related P and B frames, while B-frame errors will be isolated. Therefore, it is desirable to develop error concealment techniques to prevent error propagation from I - frames and, consequently, to improve the quality of reconstructed pictures.

5. Concluding remarks

Asynchronous transfer mode ATM is a key technology in the construction of the broadband ISDN. In the ATM network, encoded video signals are divided into short fixed length packets called cells, and can be transferred efficiently maintaning high quality by the statistical multiplexing of cells. However, in the congestion state, an appropriate congestion control is needed to avoid deterioration of transfer quality due to the loss of cells. Namely, in order to limit this loss and minimize its effect on video quality degradation, layered coding technique can be recommended. Picture elements pixels are divided in two groups in a video coding method considering cell losses in ATM-based networks. The first group is called higher priority pixels, while the second one is with lower priority pixels. The coded data of the first group are transmitted in lower priority cells. In case of network congestion, the proposed method prevents significant degradation in image quality by selectively discarding lower priority cells.

Efficient handling of video services will be one of the key factors for the successful commercialization of future broadband integrated networks. With advance in transmission as well as switching technologies, digital signal processing techniques and VLSI tecnology, increasing digital video services ranging from videophone to super - HDTV are likely to be provided. Without any bandwidth-reduction techniques for video transmission, the multiplexed traffic from subscriber access lines could easily overload the high-speed BISDN channels.

AD-HDTV provides important interoperability features at every layer of its architectures. It also provides a broad scope of features and services, allowing opportunities to tailor services to specific markets. With interoperability scope of services and extensibility characteristics, AD-HDTV will be an important part of the world's communications infractructure far into the 21 st century. AD-HDTV MPEG + + compression simultaneously provides high-quality HDTV pictures and forms the basis of AD - HDTV's reliable and robust performance.

REFERENCES

1. D. L. Mc Laren and D.T. Nguyen: *Removal of subjective redundancy from DCT-coded images.* IEE Proceedings-Part I, Vol. 138, No.5, October 1991, pp 345-350

2. Y. Sato and K. J. Sato: *Virtual path and link capacity design for ATM networks.* IEEE Journal of Selected Areas in Communications, Vol. 9, No.1, January 1991, pp 104-111.

3. M. De Prycker: *Asynchronous transfer mode - solution for broadband ISDN.* Elis Horwood, London, 1990.

4. ISO-IEC JTC1/SC2/WG11: *MPEG Document 92/No. 245.* Test Model 2, July 1992.

5. T. Kinoshita and T. Nakahaski: *Experimental HDTV codes with ATM cell loss compensation for B-ISDN.* Electronics Letters, Vol. 27, No.20, September 1991, pp 1830-1831.

6. D. J. Le Gall: *The MPEG video compression algorithm - a review.* Proceedings SPIE: Image Processing Algorithms and Techniques II, Vol. 1452, February-March 1991, San Jose, California, pp 444-457.