THE IMPACTS OF ERRORS AND DELAYS ON THE PERFORMANCE OF MPEG2 VIDEO COMMUNICATIONS

A. Mehaoua [1], *R. Boutaba* [2]

[1] University of Cambridge
Center for Communications System Research
10 Downing Street CB2 3DS Cambridge, UK
a.mehaoua@ccsr.cam.ac.uk

[2] University of Toronto, ECE Dept.
10 King's College Road
Toronto, Ontario, M5S 1A4, CANADA
rboutaba@nal.utoronto.ca

ABSTRACT

Transmission of MPEG2-encoded video is one of the most demanding applications in terms of network resources and QoS requirement. It needs high bandwidth with stringent transmission delays. It can not tolerate large variations on delays and it requires low error and loss data rates. Therefore, in order to design efficient integrated video communication systems over ATM networks, we propose in this paper to analyze the effects of errors and delays on both video signal and network performance.

1. INTRODUCTION

MPEG2 and ATM have been adopted as the key technologies for the deployment of broadcast and interactive video services. The aim of these two international standards is to provide all the advantages of transmitting variable bit rate video over packet networks, i.e. better video quality, less delay, more connections, and lower cost. However asynchronous transfer of video requires careful integration between the network and the video end systems. Numerous considerations and tradeoffs have to be considered when developing such demanding services over lossy networks. Therefore, this article analyzes the effects of errors and delays on both video quality and network performance. Alternatives and methods for dealing with these errors and delay are also discussed.

The paper is organized as follows. Section 2 addresses the effects of network errors on the video quality. In sections 3 and 4, we respectively focus on the delay and jitter impacts on the video signal, end systems and network performance. Finally, we conclude and give some recommendations for improving error resilience.

2. EFFECTS OF BIT ERRORS ON MPEG2 VIDEO QUALITY

The sources of bit errors during transmission of MPEG2 video streams over ATM networks are mainly twofold: random bit errors at the physical transmission layer, and ATM cell loss at the ATM layer due to switch buffer overflows.

The effect of a bit error in an MPEG2 video stream strongly depends on where it occurs in the bit stream syntax, or in other words in the data structure hierarchy (i.e. Sequence, Group of Pictures, Pictures, Slice, Macroblock, and Block). These errors lead to spatial and temporal error propagation in the sequence.

Table 1 from [1] and [2] gives the Bit Error Rate (BER) and the Cell Loss Ratio (CLR) required for various audiovisual services.

Before errors or losses can affect the application, they are handled by the *ATM adaptation layer* (AAL). Thus, their effects on the video quality are dependent on the behavior of the used AAL. If the selected AAL supports forward error detection (e.g. AAL1 and AAL2), the tolerance to the errors and loss is higher as shown in the Tables 1. If the AAL supports only error detection (e.g. AAL5), it may discard the entire AAL *protocol data unit* (PDU) if any error is occurring. Discarding the entire AAL PDU could actually amplify the effect of the error if the PDU size is large [3]. Figure 1 illustrates the alternative options to encapsulate the MPEG2 video packets.

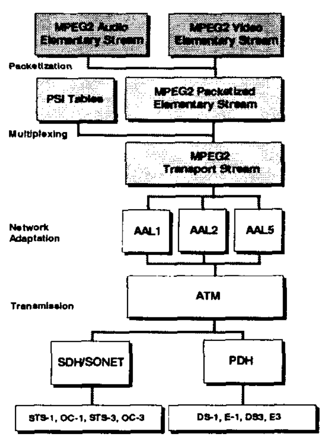

Figure 1 - Mapping MPEG2 Transport Stream Packets over ATM Networks

In 1995, the AMS[1] Technical Committee of the ATM Forum recommended that MPEG2-encoded Video-on-Demand (VOD) should be transmitted using CBR and AAL5 [4]. They also recommended the encapsulation of exactly two MPEG2 Transport Stream (TS) packets per AAL5 PDU as default but permitted the use of larger numbers.

[1] Audiovisual Multimedia Services

In [5], the authors investigate the performance of encapsulating various numbers of TS packets per AAL5 PDU in an ATM LAN environment. 10-12 packets per AAL5 PDU is found to optimize the network throughput but an additional 8-byte CS-PDU trailer has to be inserted every 2 MPEG2 TS packets within the AAL5 PDU to improve the error control capability of the end-systems.

However, AAL5 does not have facilities for FEC and jitter control and, as defined in [6], seems to be inappropriate for real-time video communications. A possible improvement is to introduce new features by means of an Audio-Visual Service Specific Convergence Sublayer (AV-SSCS) as those proposed in [7] and [8].

Some authors have tested the effect of passing corrupted AAL5 PDU to the decoder [9]. They considered an MPEG-2 video sequence of 5 Mbps and an ATM LAN with a mean bit error of 10^{-7}. Bit errors were applied at the cell level using a uniform distribution, where 79 percent of the errors fell in video information, 17 percent in audio information, and 4 percent in system information. They conclude that ignoring the CRC and passing corrupted data to the decoder reduces both the number and persistence of visible artifacts in the sequence. However, in the case of cell loss, decoding incorrect-length AAL PDUs noticeably degrades the video quality. Screen blanking, picture freezing and jitter effects like color blurring appear because the decoder could not maintain TS packet boundaries and PCR synchronization. As previously proposed in [10], the authors recommend introducing byte padding to preserve TS packet spacing and thereby picture quality.

According to their experiments, the percentage of noticeable video errors equals 58 percent when incorrect-length AAL5 PDUs are simply discarded, and equals 64 percent and 51 percent when the AAL PDU are passed to the decoder respectively without and with padding.

3. EFFECTS OF DELAYS ON MPEG2 VIDEO AND NETWORK PERFORMANCE

A number of factors contribute to the overall delay in a video communication. We briefly list them in the following:

- Encoding and decoding delays
- Picture reordering at both ends
- Encoder and decoder buffering (not including buffers for network adaptation)
- Interface delays for network adaptation between network and codec (including packetization delays)
- ATM network delays (including transmission, propagation and processing)

Table 1 also summarizes the maximum cell transfer delay and the jitter tolerance for various audio-visual services.

There is a misconception that all video services must receive quality guarantees in terms of strictly small delay bounds from the network.

If it is a one-way service (e.g. home shopping, video-on-demand), then there is flexibility on the limit for the end-to-end *cell transfer delay* (CTD). It is also accepted that a delay of several seconds occurs between the user request and the start of displaying the video sequence. This presentation delay is mainly due to the storage of two-three pictures by the destination for decoding purpose [11].

If the service is bi-directional then the CTD becomes an important factor in the presentation of real-time video. With a delay less than 150 ms, there is little impact, but serious degradation in quality can occur if the delay is more than 400 ms [2].

An important component contributes to the overall delay in a MPEG2 video communication: the coding delay. This processing delay is generally larger than the packetization and other network delays [12].

Interframe encoding requires buffering at least 2 or 3 pictures at the source and destination in order to encode and recover P- and B-pictures. For example, if three B-pictures are encoded between pair of I and/or P pictures, then a total of five pictures must be stored (i.e. I, B, B, B, P) before the B pictures can be encoded. This introduces a delay of four pictures (about 160 ms) at the source.

At the destination, the same latency should be also experienced. But to minimize it, the MPEG2 standard recommends reordering the compressed pictures prior to transmission. Each set of B pictures is transmitted after the associated reference pictures (i.e. pair of I/P pictures). This allows a delay of only one picture (about 30 ms) at the decoder before the B-pictures can be decoded and displayed.

For conversational real-time video communications, this latency is still too high since it contributes to half of the tolerated end-to-end delay. Therefore, some authors suggest to simply eliminating the use of B-pictures at the expense of higher bit rates but better video quality [13].

Conversely with one-way video services, some experiments demonstrate the advantage of increasing the end-to-end delay for the purpose of smoothing the video traffic to improve the multiplexing performance [14]. When the end-to-end delay increases to up to two-frame times, the SGM increases significantly to 35 percent. Further increases of the delay beyond three or four frames yield marginal improvements (only 2.6 percent). Other empirical measurements of MPEG over TCP and ATM recorded only 7 percent of TCP packets encountering delays of more than three MPEG frame times [15].

Thus, for a given picture quality, the designer can choose to optimize either the end-to-end-delay or the multiplexing gain. Improvements in one will degrade the performance of the other. This is another compromise to deal with.

4. EFFECTS OF JITTER ON MPEG2 VIDEO QUALITY AND DECODER MEMORY SIZE

Evaluation and control of the variations of the end-to-end delay, also called jitter, is a critical issue for video communications. Extreme jitter can have a significant impact on both the video quality and the hardware requirements to accommodate this jitter (i.e. decoder buffer size). A lengthy discussion of sources of jitter and ways to estimate it in ATM networks is provided in appendix A of the ATM Forum Video on Demand specification 1.1 [4].

Using MPEG2 System multiplex, high jitter can result in loss of synchronization of the encoder and the decoder clocks. This clock deregulation at the decoder side can manifest itself as color distortions, loss of audio/video synchronization. In extreme cases, screen freezing and blanking due to buffer overflow or underflow is also observed [9].

According to some studies, color distortion is the predominant effect exhibited by periodic jitter applied to a MPEG2 TS stream at a frequency from 0.05 to 2 Hz [9]. The jitter propagates to the Program Clock References (PCR), which are used by the decoder to recover its internal system clock. This clock must be very stable, since besides being used for controlling timing in the decoder, it is also used to derive the timing for the analog video signal (e.g. NTSC, PAL), which is the ultimate output to the viewer's television screen. The stability requirements for the various components of these analog video signals are very stringent, particularly for the color subcarrier [16]. Therefore, jitter in the PCRs will simply lead to frequency shifts in the color subcarrier and thereby to noticeable color distortion in the displayed video.

As jitter amplitudes increase in the TS, from 10 ms to 10 seconds, more severe errors are observed in the decoded video, like screen blanking and freezing [9]. These are explained by destination buffer overflow and underflow. According to the authors, less than 20 Kbytes of memory can prevent overflow and underflow of the video buffers when a 6 Mbps MPEG2 TS experiences ± 10 ms of jitter. This amount of memory is insignificant when compared with the memory requirements of the three-frame buffer (approximately 1.5 Mbps) used during the decoding process. This implies that a decoder's jitter tolerance will be constrained not by the amount of memory in the system but by the design of the phase-lock-loop (PLL).

A synchronization problem between the video and the associated audio stream is also observed when excessive jitter is measured at the destination. This artifact is referred to *skew* or *lip-sync* in [2], and it is defined as the difference in the presentation time of the video stream and the audio stream. Skew objective for a videoconferencing should be less than 20 ms if the audio is in advance of the video and less than 120 ms if the video is in advance of the audio.

There are two ways to control the jitter and minimize its effects on the video end-system. First, it is possible to limit the delay variations in the network by cell buffering and traffic smoothing, second, by absorbing these variations in the end system (i.e. in the AAL, in the MPEG2 System layer or in both).

Jitter control in the network comes with the tradeoff of having either a high minimum delay or more complex scheduling. Besides, buffering data in the network requires to carefully determine the buffer size to find the proper compromise between delay and loss. Since not all applications require jitter-free service, jitter removal might be also done outside the network when needed.

AALs can both remove delay variations (i.e. AAL1 and AAL2) and introduce delay variation of their own (i.e. AAL5). None of the currently standardized AALs is appropriate for the transmission of unconstrained real-time VBR MPEG2 streams [17].

AAL1 is dedicated to CBR video and is considered to introduce too many overheads. In addition, AAL1 provides a timestamp, the Synchronous Residual Time Stamp (SRTS), which is often redundant to the MPEG2 Program Clock Reference (PCR).

AAL2 is intended to mobile communications with a maximum transmission unit of 64 bytes [18]. For that reason it is not applicable with the 188-byte length MPEG2 TS packets.

Thus, AAL5 seems to be the best candidate. It has a lower overhead and permits the transmission of variable length packets. Because of its ubiquitous use in data applications and its application in signaling, it has been already be adopted by the ATM Forum [4] and the Digital Audio-Visual Council (DAVIC) [19] for the design of video communication systems. However, AAL5 has limited capabilities and has to be enhanced by the introduction of an audiovisual SSCS with FEC, jitter removal, and data flow multiplexing capability [17]. The design of such AV-SSCS is still an open issue.

Finally, the solution could be to limit the jitter across the network to a range that is manageable by the destination codec, and then use MPEG2 System layer features to remove it. Indeed, in case of the introduction of a small jitter at the MPEG2 System layer, i.e. less than ±4 ms, the decoders that conform to the MPEG2 standard and rely on an enhanced AAL5 should be able to dejitter the signal using their internal buffer and system clock reference.

5. CONCLUSION

In this paper, we emphasized the effects of bit errors, cell losses and delay variations on the transmitted video quality and the ATM network performance. The essential role of the ATM Adaptation Layer on reducing these adverse effects is also discussed. From this study, we are able to suggest some practical solutions to improve error resilience during transmission of MPEG-encoded video streams on error-prone networks. The increase of the occurrence of Intra-coded picture in the video sequence and the variation of the number of video slices per picture are good strategies. The reduction of the number of MPEG2 Transport Stream (TS) packets per AAL5 CS-PDU and the design of additional forward error detection and correction mechanisms at the AAL are recommended.

Finally the prioritization of the pictures before transmission have also to be considered.

6. REFERENCES

[1] M. Schwartz and D. Beaumont, 'Quality of Service Requirements for Audio-Visual Multimedia Services', ATM Forum, af-saa-94-0640, July 1994.

[2] R. O. Onvural, 'Asynchronous Transfer Mode Networks: Performance Issues', Artech House, Boston, 1994, pp.44-82.

[3] S. Dixit and P. Skilly, 'MPEG2 over ATM for Video Dial Tone Networks: Issues and Strategies', IEEE Network, Vol.9, No.5, Sept./Oct. 1995, pp.30-40.

[4] ATM Forum, 'Audiovisual Multimedia Services: Video on Demand Specification 1.1,' af-saa-0049-001, March 1997.

[5] M. Lin, D. Singe, and A. Periyannam, 'Supporting CBR-encoded MPEG2 Transport over Local ATM Networks', Multimedia Systems Journal, 1996, n.2, pp. 87-98.

[6] ITU-TS, 'Recommandation I.363 B-ISDN ATM Adaptation Layer (AAL) Specification', Geneva, March 1993, pp.38-57.

[7] K. Kanai, R. Grueter, K. Tsunoda, T. Saito and H. Esaki, 'Forward Error Correction Control on AAL5 : FEC-SSCS', IEEE ICC '96, Dallas, USA, June 1996, pp.384-391.

[8] Mehaoua, R. Boutaba and G. Pujolle, 'Proposal of an Extended AAL5 for VBR MPEG-2 over ATM', IEEE Multimedia Systems '97, Ottawa, Canada, June 1997.

[9] S. Gringeri, B. Khasnabish and al., 'Transmission of MPEG2 Video Streams over ATM', IEEE Multimedia, Vol.5, No.1, January 1998, pp.58-70.

[10] A. Mehaoua and R. Boutaba, 'Performance analysis of cell discarding techniques for best effort video communications over ATM networks', in Computer Networks And ISDN Systems Journal, vol.29, issue 17-18, February 1998, pp. 2021-2037.

[11] G. Karlsson, "Asynchronous Transfer of Video," IEEE Communications Magazine, Vol.34, No.8, August 1996, pp. 118-126.

[12] N. Ohta, 'Packet Video: Modelling and Signal Processing', Artech House, Boston, 1994.

[13] Cisco Systems, 'White Paper: Video over ATM and Existing Networks', November 1995, pp.1-26. available at http://www.cisco.com

[14] T.V Lakshman, A. Ortega and A.R. Reibman, 'VBR Video: Tradeoffs and Potentials', Proceedings of the IEEE, Vol.86, No.5, May 1998, pp.952-973.

[15] R.P. Tsang, D.H.C. Du, and A. Pavan, 'Experiment with Video Transmission over ATM Network', Multimedia Systems Journal, 1996, Vol.4, No.4, pp. 151-171.

[16] D. Minoli, 'Video Dialtone Technology', McGraw-Hill, New-York, 1995.

[17] X. G. Adanez, O. Verscheure and P. Frossard, 'Reliable Transmission of MPEG-2 VBR Video Streams over New Network Adaptation and ATM Adaptation Layers', IEEE ATM Workshop 97, Lisbon, Portugal, May 1997.

[18] ITU-TS , 'recommandation I.363.2 : B-ISDN ATM Adaptation Layer Type 2 Specification', Madrid, November 1996.

[19] DAVIC, 'Digital Audio-Visual Council specification 1.3 part 7: High and Mid Layer Protocols', Technical specification revision 6.3, May 1998.

Service	Bit Rate	QoS required	BER 1	BER 2	CLR 1	CLR 2	Delay	Jitter
Two-way Video Communication								
Videophone	64 Kbps - 2 Mbps CBR (H.261)	30 min error free	10^{-7}	-	10^{-8}	8×10^{-6}	200-500 ms	130 ms
Videophone	2 Mbps VBR (MPEG2)	30 min error free	3×10^{-11}	1.2×10^{-7}	10^{-8}	8×10^{-6}	300 ms	-
VideoConference	5 Mbps VBR (MPEG2)	30 min error free	10^{-11}	8×10^{-8}	4×10^{-9}	5×10^{-6}	300 ms	-
Video Distribution								
VCR quality	1.5 Mbps CBR (MPEG1)	20 min error free	4×10^{-11}	1.4×10^{-7}	10^{-8}	9.5×10^{-6}	5 sec.	6.5 ms
TV quality	10 Mbps VBR (MPEG2)	30 min error free	6×10^{-12}	5.4×10^{-8}	2×10^{-9}	4×10^{-6}	1 sec.	1 ms
Studio TV quality	15 Mbps CBR (MPEG2)	1 hour error free	2×10^{-12}	1.5×10^{-7}	10^{-9}	10^{-8}	1 sec.	1 ms
HDTV quality	25-80 Mbps VBR (MPEG2)	2 hour error free	3×10^{-13}	1.2×10^{-8}	10^{-10}	8×10^{-7}	1 sec.	0.8 ms