# Joint Impact of MPEG-2 Encoding Rate and ATM Cell Losses on Video Quality

Olivier Verscheure, Pascal Frossard* and Maher Hamdi
Institute for computer Communications and Applications
Swiss Federal Institute of Technology
Email: Olivier.Verscheure@epfl.ch
URL: http://icawww.epfl.ch/

## Abstract

*This work addresses the analysis of the end-user perception of CBR MPEG-2 video services transmitted over ATM/AAL5 networks. We first independently study how the perceptual video quality behaves with the constant encoding bit rate (CBR) and the network cell loss ratio (CLR). We then jointly analyze the impact of these two components onto the user-oriented QoS. All the quality assessments are based on a perceptual quality metric, MPQM, which proved to behave consistently with human judgements [1]. We show that, when considering video transmission over lossy networks, there is an optimal bit rate to be determined that maximizes the end-user perception of the service under some given network conditions.*

## 1 Introduction

The Asynchronous Transfer Mode (ATM) technology is reaching a certain level of maturity that permits its deployment in local as well as in wide area networks. One of the key points of such a technology is that it may provide statistical guarantees on performance. In other words, each user may agree with the network on a set of performance parameters (i.e. Quality of Service (QoS) parameters) which will have to be met.

Audiovisual applications (e.g. video conferencing, video on demand, teleteaching, etc.) are foreseen as one of the major users of such broadband networks. At the heart of this revolution is the digital compression of audio and video signals. The biggest advantage of compression resides in data rate reduction which results in reduction of transmission costs. The choice of the compression algorithm mostly depends on the available bandwidth or storage capacity and the features required by the application. The MPEG-2 standard [2], a truly integrated audio-visual standard developed by the International Organization for Standards (ISO), is capable of compressing NTSC or PAL video into an average bit rate of 3 to 6 Mbits/s with a quality comparable to analog CATV [3]. However, a lot of work remains to be done to optimize these audiovisual applications so that they can be offered at attractive prices. In other words, the user expects an adequate audiovisual quality at the lowest possible cost. In the case of video transmission over packet networks, the end-to-end User-oriented Quality of Service (U-QoS) results both from the video encoding quality and the degradations due to data loss, delay and delay jitter during the transmission.

This work focuses on the analysis of the quality perceived by the end-user in the case of CBR MPEG-2 video transmission over ATM networks. The paper is organized as follows: Section 2 first introduces the MPEG-2 video and system standards and then briefly describes the impact of data loss onto the reconstructed video sequence. The analysis of the user-oriented QoS is subject of the Sec. 3. This section first starts with some preliminary studies as well as the description of the experimental setup and then presents some important results concerning the impact of both data loss and encoding bit rate on the user perception of the service. Concluding remarks are given in Sec. 4 .

*now with the Signal Processing Laboratory (LTS)

0-7803-4984-9/98/$10.00 ©1998 IEEE.

## 2 MPEG-2 over Packet Networks

### 2.1 MPEG-2 Backgrounder

An MPEG-2 video stream is highly hierarchically structured as illustrated in Fig. 1. The stream consists of a sequence composed of several pictures. The MPEG-2 video standard defines three different types of pictures : intra-coded (I-), predicted (P-) and bidirectional (B-) pictures. The use of these three picture types allows MPEG-2 to be robust (I-pictures provide error propagation reset points) and efficient (B- and P-pictures allow a good overall compression ratio). Each picture is composed of slices which are, by definition, a series of macroblocks. Each macroblock ($16 \times 16$ pixels) contains 4 blocks ($8 \times 8$ pixels) of luminance and 2, 4 or 8 blocks of chrominance depending on the chroma format.

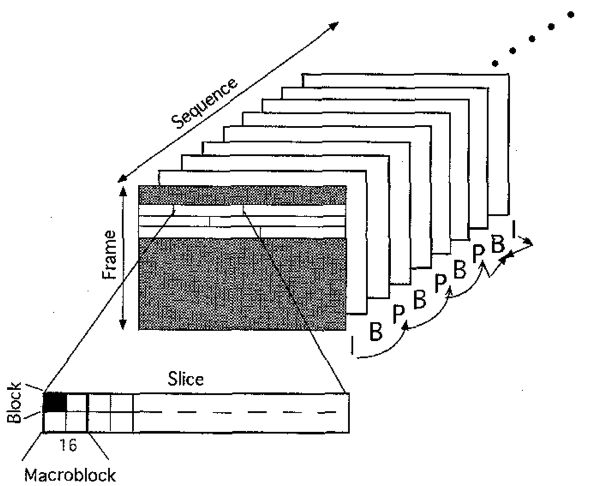


**Figure 1. MPEG-2 video structure**

Before being transmitted, such a video stream goes through the MPEG-2 Transport Stream (TS) layer. Basically, the stream is first segmented into variable-length Packetized Elementary Stream packets and then subdivided into fixed-length TS packets of 188 bytes. Moreover, it is worth noting that a non-encoded header (i.e. syntactic information) is inserted before each of the following information elements: sequence, Group of Pictures (GoP), picture, slice, TS and PES. In general, when a header is damaged, the underlying information is lost.

### 2.2 MPEG-2 Sensitivity to Data Loss

In an MPEG-2 video stream, data loss reduces quality depending strongly on the type of the lost information. Losses in syntactic data, such as headers and system information, affect the quality differently than losses of semantic data such as pure video information (e.g. motion vectors, DCT coefficients, etc.). Furthermore, the quality reduction depends also on the location of the lost semantic data due, for instance, to the predictive structure of MPEG-2 video coded streams.

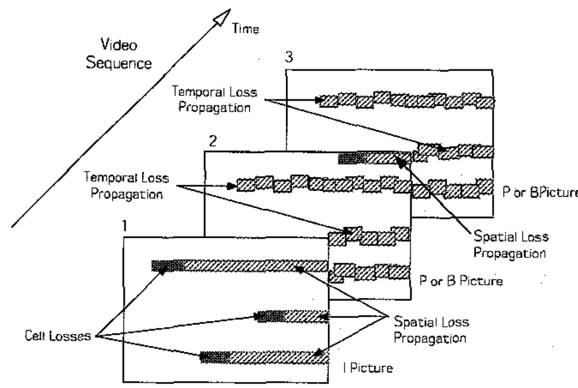


**Figure 2. Data loss propagation in MPEG-2 video streams.**

Figure 2 shows how network losses map onto visual information losses in different types of pictures. Indeed, data loss spreads within a single picture up to the next resynchronization point (e.g. picture or slice headers) mainly due to the use of differential coding, run-length coding and variable length coding. This is referred to as spatial propagation and, it may damage any type of picture. When loss occurs in a reference picture (intra-coded or predictive frame), the damaged macroblocks will affect the non intra-coded macroblocks in subsequent frame(s) which reference the errored macroblocks. This is known as temporal propagation and is due to inter-frame predictions.

Error concealment algorithms have already shown that it is possible to reduce the impact of data loss on the visual information [4, 5, 6]. These error concealment algorithms include, for example, spatial interpolation, temporal interpolation and early resynchronization techniques. Early resynchronization decoding

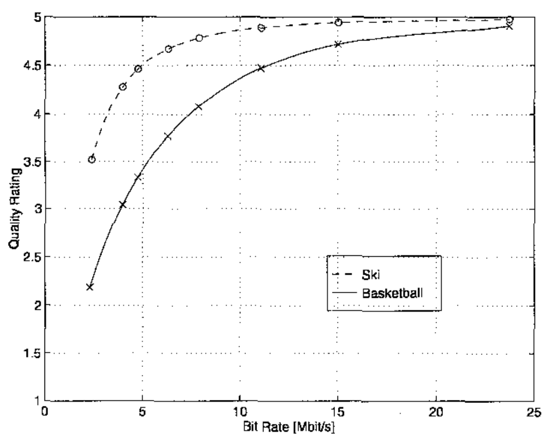


**Figure 3. MPQM versus encoding bit rate.**

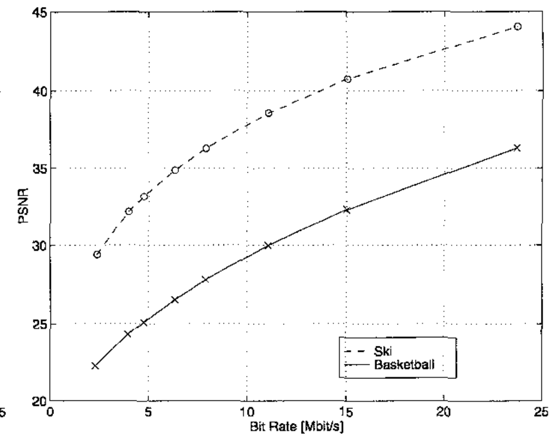


**Figure 4. PSNR versus encoding bit rate.**

techniques limit the spatial propagation of errors by decoding some semantic information that is normally discarded from the damaged MPEG-2 video streams. All these techniques only allow to increase the range of acceptable cell loss ratios. Data loss may still involve very annoying perturbation in the decoded video, especially when frames are lost.

## 3 User-oriented QoS Analysis

As stated before, the quality of a video service perceived by an end-user results both from the video encoding quality and the degradations due to data loss, delay and delay jitter during the transmission. In this work, we analyze the impact of video encoding and data loss only on the end-to-end U-QoS in the case of CBR MPEG-2 video services delivery.

### 3.1 Encoding Quality

We first analyse how the video quality behaves with the constant encoding bit rate. Several studies have shown that a good estimation of subjective video quality has to incorporate some modeling of the Human Visual System [7]. In this work, we used a computational quality metric for moving pictures, namely the MPQM metric, which proved to behave consistently with human judgements [1]. The resulting quality rating is a value between 1 and 5 according to the ITU-R 500.3 standard [8] (the higher the value, the better the quality).

In our simulations, two 128 frame-long sequences conforming to the ITU-T 601 format have been used (i.e. Basketball and Ski). These sequences have been encoded with a constant bit rate (CBR) as interlaced video, with a structure of 12 images per GOP, 2 B-pictures between every reference picture and one slice per line. For that purpose, the TM5 MPEG-2 software encoder has been used [9].

Figure 3 shows how the video quality behaves according to the constant encoding bit rate for the Ski and BasketBall sequences. An important result that can be extracted from the graph is that the perceptual quality saturates at high bit rates. Increasing the bit rate may thus result, at some point, in a waste of bandwidth since the end user does not perceive an improvement in quality anymore. Additional compressed video sequences show results with the same behavior in [1]. Such a quality saturation is however not captured by the PSNR[1] metric as illustrated in Fig. 4. Many rate control algorithms are using the PSNR as the performance evaluation tool. Both figures show that even a gain of several dBs in PSNR may not reflect any change from the user viewpoint.

[1]PSNR stands for Peak Signal-to-Noise Ratio

### 3.2 Cell Loss Impact on Quality

Up to this point, we did not consider any data loss in the bitstream. We now study the practical case of CBR MPEG-2 transmission over ATM/AAL5 networks [10].

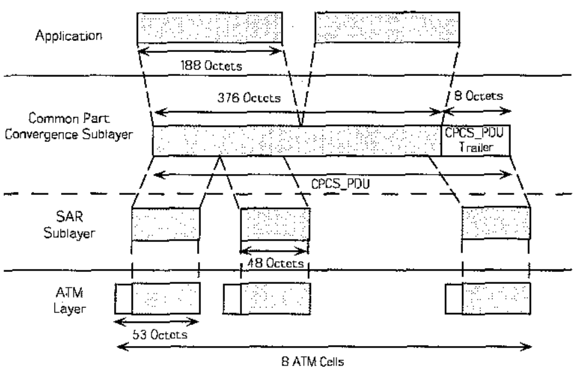


**Figure 5. AAL5 Segmentation Mechanism.**

Before being transmitted, the MPEG-2 video bitstream is encapsulated into 18800-bytes length Packetized Elementary Stream (PES) packets and divided into fixed length Transport Stream (TS) packets by the MPEG-2 system encoder.

The transmission of video over AAL5 is based on the approved ATM Forum Video on Demand specification [11]. This scheme packetizes two TS packets into a single AAL5-SDU. The AAL5 adds its 8-byte trailer and the resulting AAL5-PDU is segmented into 8 ATM cells without any padding (see Fig. 5).

The simulation framework is composed of two multimedia workstations and one ATM switch. The switching module is implemented as a multiplexer with limited buffer size. To generate cell losses in the multimedia streams, we load the multiplexing stage with background traffic provided by several On-Off sources. This type of source model is widely used to simulate a multiplex of traffic such as the one that could be found at the entrance of an ATM switch.

The decoder provides the error concealment techniques briefly described in Sec. 2.2 and tested in [4]. These techniques have been used for different reasons. The first is to be consistent with real implementations and, the second, to be able to perform the perceptual measurements. Indeed, the human visual system models currently developed and the derived metrics have been tested for errors below the so-called *supra-threshold* of vision. The problem is that in general, the errors due to cell losses generate highly visible artifacts (i.e. holes) in the reconstructed sequence. By using concealment techniques, most of the artifacts may be considered as being below the supra-threshold of vision, making the perceptual measure accurate[2].

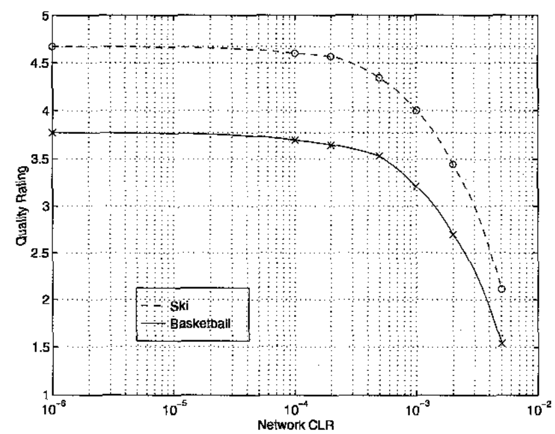


**Figure 6. MPQM versus Network CLR for a fixed CBR=6 MBits/s.**

Figure 6 presents how the video quality behaves with the Cell Loss Ratio (CLR) (i.e. background load). It is shown that, on a semi-logarithmic scale and for a given average bit rate, first the video quality remains constant with the CLR. This constant value corresponds to the encoding quality. Then, beyond a certain CLR, the perceptual quality quickly drops.

### 3.3 End-User Perception of the Service

In the last two subsections, we independently analyzed the impact of both the encoding bit rate and the network CLR on the end-user perception of the video service. We will show now that these two components are actually not independent resulting in a complex relationship gathering all the parameters together. For this purpose, we used the same simulation setup with the "Ski" sequence only.

[2] Above the suprathreshold of vision, the MPQM metric partly looses its accuracy

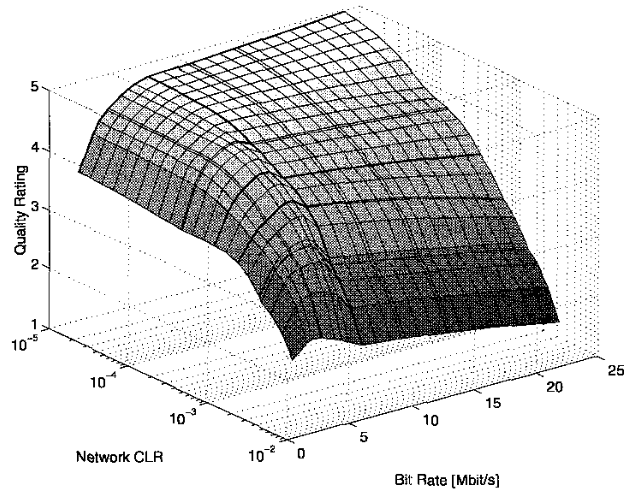


**Figure 7. MPQM versus encoding bit rate and network CLR.**

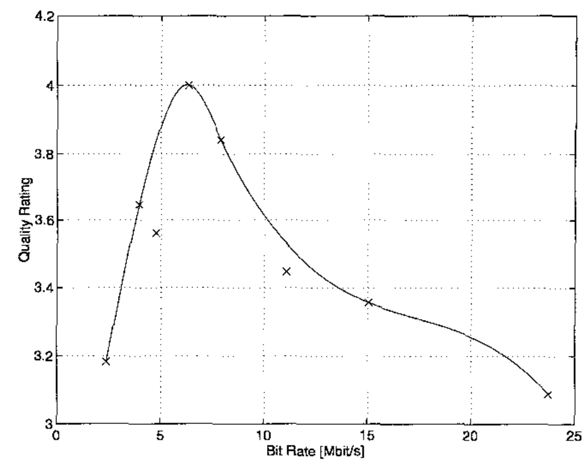


**Figure 8. MPQM versus encoding bit rate for a fixed CLR=$10^{-3}$.**

Figure 7 shows how the user-oriented video quality varies with both the encoding bit rate and the network cell loss ratio.

One of the most interesting feature to be extracted from the graph is the evolution of the quality with the encoding bit rate for a non-zero network CLR. Indeed, the video quality first increases up to a certain bit rate and then smoothly decreases. This behavior is better illustrated on Fig. 8 which has been extracted from the previous graph.

Such an important behavior may be explained by the following: since the CLR is equal to the ratio between the cell loss rate (number of lost cells per second) and the cell rate (i.e. bit rate), then the lower the bit rate, the lower the cell loss rate for a fixed CLR. Also, the lower the cell loss rate, the lower the number of lost cells per frame on average. Therefore, the lower the bit rate, the higher the CLR for an equivalent perceived degradation (see Fig. 9).

Therefore, at a given threshold, a bit rate increment involves a smaller encoding quality improvement than a quality degradation due to video data loss.

Finally, when considering video transmission over lossy networks, not only it is bandwidth consuming to increase the encoding bit rate above a certain threshold (which varies according to the scene complexity), it is also quality consuming. In other words, when the user-oriented QoS is not high enough, an increase of the encoding bit rate at a constant CLR may even degrade the quality depending on the position of the working point on the 3D graph presented herebefore. There is actually an optimal bit rate to be determined that maximizes the end-user perception of the service under some given network conditions (i.e. network impairments).

It has to be noticed that such a conclusion is general enough to be applied to a different encoding system and/or networking technology. It may be argued that the packet loss ratio increases along with the transmission rate in some packet networks. However, such an observation enforces our analysis. Indeed, the higher the PLR, the higher the number of loss hits in average.

## 4 Conclusion and Future Works

In this paper, we jointly analyzed the impact of MPEG-2 encoding rate and ATM cell losses on the user-oriented video quality. It appeared that the encoding bit rate has to be carefully chosen as soon as cell losses occur. The video quality may indeed decrase when increasing the encoding bit rate under the same network conditions. This conclusion still holds for a different encoding system and networking technology.

Further works are being carried out. We are cur-

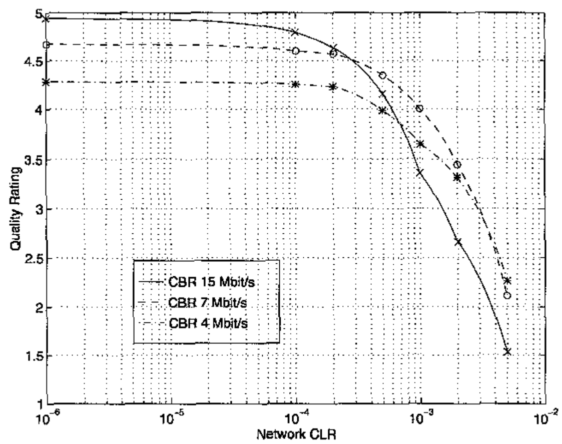


**Figure 9. MPQM versus network CLR for CBRs of 4, 7 and 15 MBits/s.**

rently working on an analytical analysis of this joint impact on the video quality [12]. We are also investigating the effect of higher order statistics of the loss process (i.e. variance, correlations) on the perceived quality. Finally, a similar set of experiments is being applied to H.263 over IP networks.

## References


[1] C. J. van den Branden Lambrecht and O. Verscheure, "Perceptual Quality Measure using a Spatio-Temporal Model of the Human Visual System", *in Proceedings of the SPIE*, vol. 2668, pp. 450–461, San Jose, CA, January 28 - February 2 1996.

[2] ISO/IEC JTC 1, *Information Technology - Generic Coding of Moving Pictures and Associated Audio Information - Part 1, 2 and 3*, ISO/IEC JTC 1, 1996.

[3] B. G. Haskell, A. Puri and A. N. Netravali, *Digital Video: an Introduction to MPEG-2*, Digital Multimedia Standards Series. Chapman and Hall, 1997.

[4] C. Lopez Fernandez, A. Basso and J.-P. Hubaux, "Error Concealment and Early Resynchronization Techniques for MPEG-2 Video Streams Damaged by Transmission over ATM Networks", *in SPIE*, San Jose, USA, February 1996.

[5] S. Lee, J. Youn and S. Jang, "Transmission error detection, resynchronization, and error concealment for mpeg-2 video decoder", *in SPIE*, 1993.

[6] S. Aign and K. Fazel, "Temporal and spatial error concealment techniques for hierarchical mpeg-2 video coded", *in ICC'95*, 1995.

[7] C. J. van den Branden Lambrecht, *Perceptual Models and Architectures for Video Coding Applications*, PhD thesis, Swiss Federal Institute of Technology, CH 1015 Lausanne, Switzerland, 1996.

[8] ITU-R Commity, "ITU-R 500.3 Recommendations and Reports", *in ITU*, vol. XI. 1986.

[9] C. Fogg, "mpeg2encode/mpeg2decode", *in MPEG Software Simulation Group*, 1996.

[10] X. Garcia Adanez, O. Verscheure and J.-P. Hubaux, "New Network and ATM Adaptation Layers for Real-Time Multimedia Applications: A Performance Study Based on Psychophysics", *in COST-237 workshop*, Barcelona, Spain, November 1996.

[11] The ATM Forum, "Video on Demand Specification 1.0", Jan. 1996.

[12] O. Verscheure, P. Frossard and M. Hamdi, "MPEG-2 Video Services over Packet Networks: Joint Effect of Encoding Rate and Data Loss on User-Oriented QoS", *in 8th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV)*, Cambridge, UK, July 8-10 1998.