Real-Time Imaging **5**, 305–314 (1999)
Article No. rtim.1999.0175, available online at http://www.idealibrary.com on IDEAL®

User-Oriented QoS Analysis in MPEG-2 Video Delivery

We address the problem of video quality prediction and control for high-resolution video transmitted over lossy packet networks. In packet video, the bitstream flows through several subsystems (coder, network, decoder); each of them can impair the information, either by data loss or by introducing some delay. However, each of these subsystems can be fine-tuned in order to minimize these problems and to optimize the quality of the delivered signal, taking into account the available bitrate. The assessment of the end-user quality is a non-trivial issue. We analyse how the user-perceived quality is related to the average encoding bitrate for variable bit rate MPEG-2 video. We then show why simple distortion metrics may lead to inconsistent interpretations. Furthermore, for a given coder setup, we analyse the effect of packet loss on the user-level quality. We then demonstrate that, when jointly studying the impact of coding bit rate and packet loss, the reachable quality is upperbound and exhibits one optimal coding rate for a given packet loss ratio.

Oliver Verscheure, Pascal Frossard[1] and Maber Hamdi

Institute for Computer Communications and Applications, Swiss Federal Institute of Technology, Lausanne, Switzerland

Introduction

The transmission of multimedia streams requires a network capable of handling different types of data. For several years, the solution has been considered to be the *asynchronous transfer mode* (ATM). ATM is the network technology for the broadband integrated services digital network (B-ISDN). Now the role of ATM is being challenged by the success of the Internet and other IP-based networks due to the new developments of integrated and differentiated services.

A truly integrated network will have to cope with different traffic characteristics and quality requirements in terms of delay, delay jitter and data loss. Providing integration of heterogeneous traffic and adequate QoS to users has proved difficult to achieve.

Work remains to be done to optimize multimedia applications so they can be offered at attractive prices. In other words, the user expects an adequate audio-visual quality at the lowest possible cost. From the user's viewpoint, in the case of video transmission over packet networks, both the encoding and the transmission processes affect the quality of service. The most economic offering can thus only be found by considering the entire system and not by optimization of individual system components in isolation [1, 2].

This paper is organized as follows: we first introduce the MPEG-2 video and system standards. We the briefly describe the impact of data loss on the reconstructed video sequence. Useful video quality metrics are also

[1] P. Frossard is now with the Signal Processing Laboratory of the Swiss Federal Institute of Technology.
Corresponding author: P. Frossard

1077-2014/99/100305 + 10 $30.00/0

described, particularly the MPQM which is based on a vision model. The impact of MPEG-2 rate and data loss on quality is then studied. Finally, we analyse the joint impact of MPEG-2 rate and data loss on video quality.

MPEG-2 over Packet Networks

MPEG-2 backgrounder

The choice of the compression algorithm depends on the available bandwidth or storage capacity and the features required by the application. The MPEG-2 standard [3], 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 [4].

An MPEG-2 video stream is hierarchically structured as illustrated in Figure 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 (16×16 pixels) contains four blocks (8×8 pixels) of luminance and 2, 4 or 8 blocks of chrominance depending on the chroma format. Motion estimation is performed on macroblocks while the DCT[2] is calculated on blocks. The resulting DCT coefficients are quantized and variable length coded. The quantizer comes from the multiplication of a Quantizer Scale, MQUANT, and the corresponding element of a Quantizer Matrix. In general, the higher the MQUANT value, the lower the bit rate but also the lower the quality (well-known from the rate-distorsion theory).

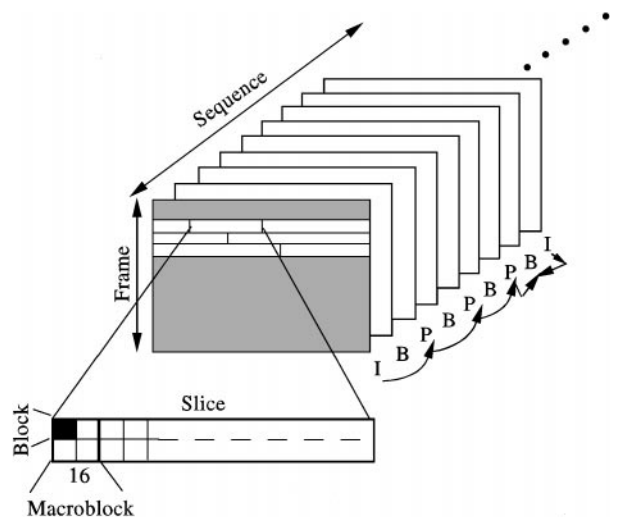

Figure 1. MPEG-2 video structure.

Before being transmitted, 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. 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.

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 of 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 on the location of the lost semantic data due, not only to the predictive structure of MPEG-2 video coded streams, but also to the visual relevance of the data.

Figure 2 illustrates how network losses map onto visual information losses in different types of pictures. 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 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.

However, the error visibility may be dramatically reduced by means of error concealment techniques.

[2] DCT stands for Discrete Cosine Transform.

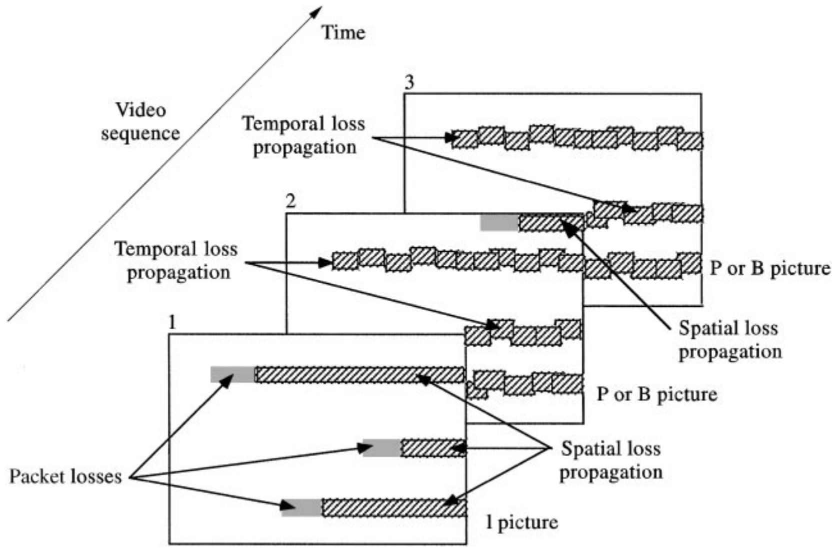

Figure 2. Data loss propagation in MPEG-2 video streams.

These error concealment algorithms include, for example, spatial interpolation, temporal interpolation and early resynchronization techniques. The MPEG-2 standard proposes an elementary error concealment algorithm based on motion compensated techniques. Mainly, it estimates the motion vectors for the lost macroblock by using the motion vectors of neighboring macroblocks in the affected picture (provided these have not also been lost). This improves the concealment in moving picture areas. However, there is an obvious problem with errors in macroblocks whose neighboring macroblocks are intra-coded, because there are ordinarily no motion vectors associated with them. To circumvent this problem, the encoding process can be extended to include motion vectors for intra macroblocks[3].

Error concealment techniques may, in general, efficiently decrease the sensitivity to data loss. However, none of these techniques is perfect. Data loss may still involve annoying degradation in the decoded video.

Video quality metrics

A quality metric often used for audiovisual signals is the peak-to-noise ratio (PSNR). Many studies have shown that this metric is poorly correlated with human perception since it does not take visual masking into consideration. In other words, every errored pixel contributes to a decrease in PSNR even if the error cannot be perceived. Recent research has therefore addressed the issue of video quality assessment by means of metrics based on the properties of the human visual system. All these metrics fall into one of the following categories: (a) metrics based on a mathematical fit of a subjective rating function obtained by intensive psychovisual experiments (e.g. $\hat{S}$ [5]) and (b) metrics relying on a model of the human visual system (e.g. JND [6], MPQM [7]). Metrics belonging to the latter category perform usually better [8].

In [7], a spatio-temporal model of human vision has been developed for the assessment of video coding quality [8, 9]. The model is based on the following properties of human vision:

- The responses of the neurons in the primary visual cortex are band limited. The human visual system has a collection of mechanisms or detectors (termed "channels") that mediate perception. A channel is characterized by a localization in spatial frequency, spatial orientation and temporal frequency. The responses of the channels are simulated by a three-dimensional filter bank.
- The channels can be considered to be independent. Perception can thus be predicted channel by channel without interaction.

[3] Some MPEG-2 encoder chips automatically produce concealment motion vectors for all intra-coded macroblocks.

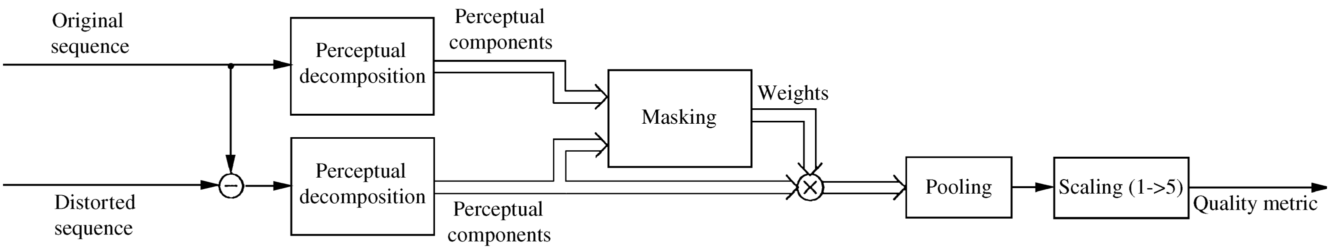

Figure 3. Moving Pictures Quality Metric (MPQM) block diagram.

- Human sensitivity to contrast is a function of both frequency and orientation. The *contrast sensitivity function* (CSF) quantizes this phenomenon by specifying the detection threshold for a stimulus as a function of frequency.
- Visual masking accounts for inter-stimuli interferences. The presence of a background stimulus modifies the perception of foreground stimulus. Masking corresponds to a modification of the detection threshold of the foreground according to the local contrast of the background.

This model has been used to build a computational quality metric for moving pictures [8] which behaves consistently with human judgments. The metric, called moving pictures quality metric (MPQM), first decomposes the original sequence and a distorted version of it into perceptual channels. A channel-based distortion measure is then computed while accounting for contrast sensitivity and masking (see Figure 3). Finally, the data is pooled over all the channels to compute the quality rating which is then scaled from 1 to 5 [10]. This quality scale is used for subjective testing in the engineering community (see Table 1).

Impact of MPEG-2 Rate and Data Loss on Quality

In this section, we first describe the experimental setup used throughout this work. We then study how the video quality behaves according to the quantizer scale factor (MQUANT) in an MPEG-2 OL-VBR[4] encoding scheme. We also analyse how the average encoding bit rate is affected by this MQUANT. We then derive a mathematical relation modeling the impact of the average variable rate of the video encoding quality.

Table 1. Quality scale generally used for subjective testing in the engineering community

Rating	Impairment	Quality
5	Imperceptible	Excellent
4	Perceptible, not annoying	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very annoying	Bad

Finally, we study how the video quality decreases when the data loss ratio is increased, for a fixed average encoding bit rate.

Experimental setup

The experimental testbed is composed of four parts (see Figure 4):

- An MPEG-2 software encoder, which is composed of an open-loop VBR TM5 video encoder [11] and a transport stream encoder. Four sequences conforming to the ITU-T 601 format were used (i.e. Football, News, Ski and Barcelona). All these sequences are very different in terms of spatial and temporal complexities. They were encoded, as interlaced video, with a structure of 12 images per GOP and two B-pictures between every reference picture in an OL-VBR mode. The following MQUANTs were used: 6, 10, 16, 20, 28, 32, 36, 40 and 48. Motion vectors were generated for all intra-coded macroblocks. It is to be noted that the OL-VBR encoding quality is not affected at all when introducing these extra motion vectors. Before being transmitted, each MPEG-2 video bitstream was encapsulated into 18 800-bytes length Packetized Elementary Stream (PES) packets and divided into fixed length Transport Stream (TS) packets by the MPEG-2 system encoder.

[4] OL-VBR stands for Open-Loop Variable Bit Rate (constant quantizer scale over the whole sequence).

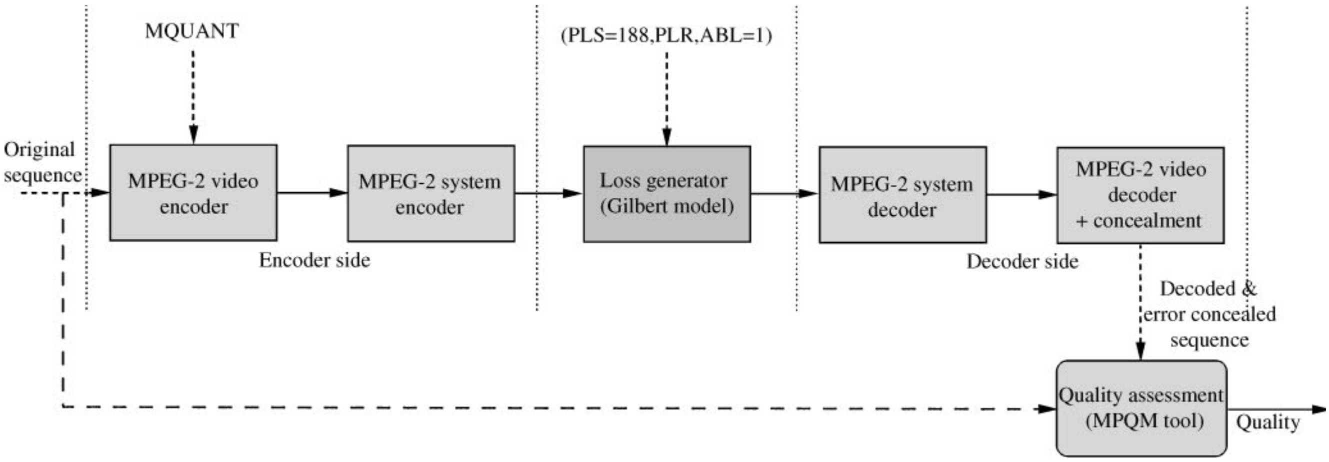

Figure 4. Experimental testbed.

- A model-based data loss generator was used to simulate packet network losses. For this purpose, we used a two-state Markovian model (Gilbert model [12], see Figure 5).

States 0 and 1 respectively correspond to a correct and an incorrect packet reception. The transition rates between the states control the length of the bursts of errors. Hence, there are three parameters to be controlled: the packet loss size (PLS), the packet loss ratio ($PLR = p/p+q$) and the average length of a burst of errors ($ABL = 1$). In our simulations, we imposed a non-bursty ($ABL = 1$) TS packets ($PLS = 188$ bytes) loss process and made the packet loss ratio vary between 10^{-2} and 10^{-7}.

- Video quality was evaluated by means of the MPQM tool presented in the previous section. The per-frame quality values given by the MPQM tool were gathered together thanks to a Minkowski summation [13] (exponent $\beta = -2$). This summation, along with the correct exponent, gives a result that is lower than the simple average quality which is too optimistic (i.e. the subjective quality evaluated over a set of frames is lower than the average of the per-frame quality values).
- The last part is an MPEG-2 software decoder constituted by both a TS decoder and a video decoder. The video decoder provides the motion compensated concealment technique briefly explained in the previous section. This technique was chosen for different reasons. The first is to be consistent with real implementations. The second is to be able to perform the perceptual measurements. Indeed, the vision model currently developed and the derived metrics have been tested for errors below what is called *suprathreshold*[5]. The problem is that, in general, the degradations due to data losses generate highly visible artefacts (i.e. holes) in the sequence and these errors are all above this suprathreshold. By using concealment techniques, most of the artifacts may be considered as being below the suprathreshold of vision, making the perceptual measure accurate.

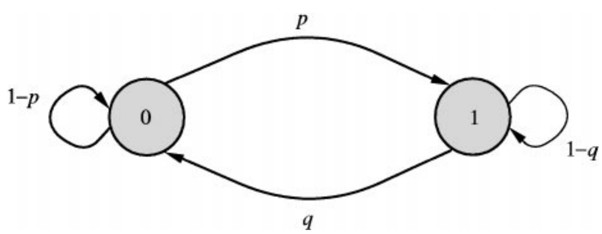

Figure 5. Two-State Makov Chain: Gilbert Model.

MPEG-2 VBR encoding impact on video quality

First, we study how the OL-VBR encoding process influences video quality on a GOP basis. Figures 6 and 7 show how the quality is affected by the MQUANT parameter using, respectively, the PSNR metric and MPQM tool to measure it. While the PSNR versus MQUANT curve may be represented by a decreasing exponential [14], it is to be noted that the MPQM metric exhibits a linear relationship with MQUANT. We

[5] Two to three times above the threshold of vision which corresponds to the threshold of visibility of the noise.

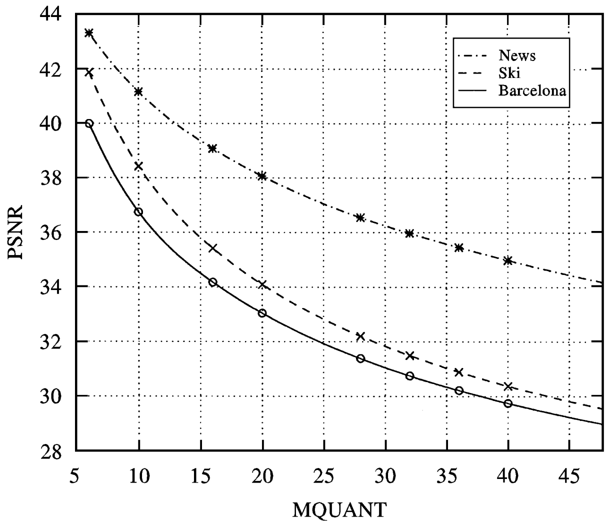

Figure 6. PSNR versus quantizer scale factor for three different scenes.

verified this important behavior for all the GOPs of the four sequences constituting our testbed. The same characteristic has recently been observed through user's subjective evaluation [15]. Computer simulation results as well as the corresponding fits are represented on Figure 7 for both the "Barcelona" and "News" sequences. The encoding quality Q_E is approximated by a function of the form:

$$Q_E = \chi_Q \times MQUANT + Q_0 \tag{1}$$

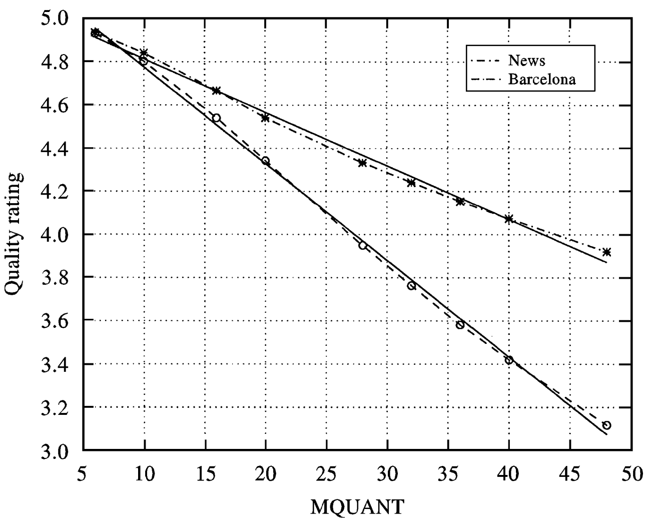

Figure 7. MPQM versus quantizer scale factor for two different scenes. Fitting parameters for Eqn (1): News: ($\chi_Q = -0.025$, $Q_0 = 5.062$), Barcelona: ($\chi_Q = -0.045$, $Q_0 = 5.226$).

where parameters χ_Q and Q_0 have been obtained by the least mean squares method (see Figure 7). The slope χ_Q is directly related to the complexity of the sequence: the higher the encoding complexity, the higher the absolute value of χ_Q. This remark may be verified on the graph. The video sequence "News" is a *Head and Shoulder* type of sequence and does not contain any high spatio-temporal complexity. The absolute value of the slope is thus smaller than for the "Barcelona" video sequence. The value of Q_0 will always be close to 5 (highest quality).

The linear relation between the video quality and quantizer scale factor may have several impacts on the design of, for instance, perceptual rate controllers or consistent quality regulators operating in real-time.

Now, we have an idea of how the encoding quality behaves according to the MQUANT. We need then to study how the average output bit rate if affected by this MQUANT. In [4], it was demonstrated that a power function curve was a good approximation to represent the relation between the quantizer scale factor and the average bit rate:

$$\bar{R} = \chi_R \times MQUANT^{-\xi_R} \tag{2}$$

in which $\bar{R}$ represents the average output bit rate and the parameters χ_R and ξ_R are related to the encoding complexity of the set of frames over which the Minkowski summation is performed (a GOP in this case).

Figure 8 illustrates this behavior very well. The parameters χ_R and ξ_R have been obtained again by minimizing the mean square error.

Finally, by combining Eqns (1) and (2), we derive a model for describing how the video quality behaves according to the average encoding bit rate:

$$Q_E = \chi_Q \times \left(\frac{\bar{R}}{\chi_R}\right)^{-\frac{1}{\xi_R}} + Q_0 \tag{3}$$

As stated before, the three main parameters χ_Q, χ_R and ξ_R are somehow related to the spatio-temporal complexity of the set of frames considered (Q_0 will always be around 5.0). However, in this work, we did not investigate this relation any further.

Computer simulation results and the corresponding fitting curve using the equation herebefore are represented in Figure 9.

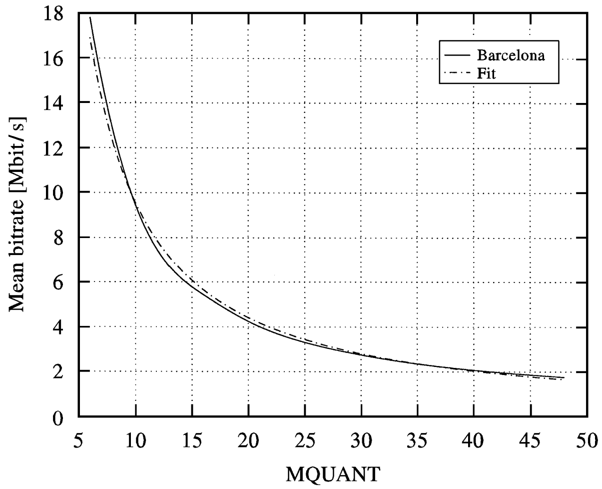

Figure 8. Average output encoding bit rate versus quantizer scale factor (MQUANT) for Barcelona. Fitting parameters for Eqn (2): ($\chi_R = 124.762$, $\xi_R = 1.116$).

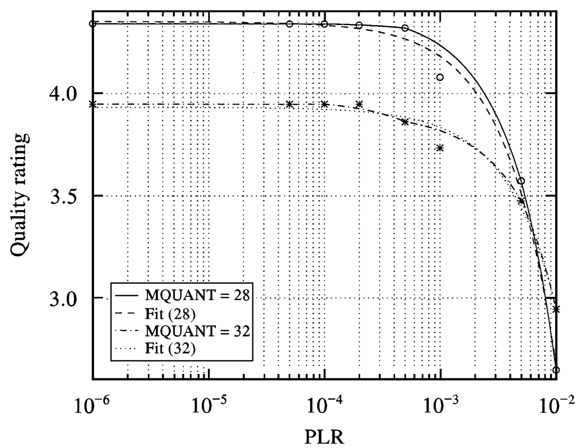

Figure 10. *MPQM* versus *PLR* ($ALB = 1$, $PLS = 188$) for $MQUANTs = \{28, 32\}$ using the Barcelona sequence. Fitting parameters for Eqn (4): $MQUANT = 28$, ($Q_E = 4.352$, $\chi_L^* = 168.162$) and for $MQUANT = 32$, ($Q_E = 3.934$, $\chi_L^* = -98.351$).

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 after a certain bit rate. However, such saturation of quality is not well captured by the PSNR.

The average bit rate after which the quality does not increase significantly may be shifted to the left by means, for instance, of an adaptive quantization scheme [16]. Indeed, adaptive quantization aims at spatially uniformizing the coding noise by adjusting the *MQUANT value* on a macroblock basis. Therefore, the same perceptual quality may be reached at a lower average bit rate.

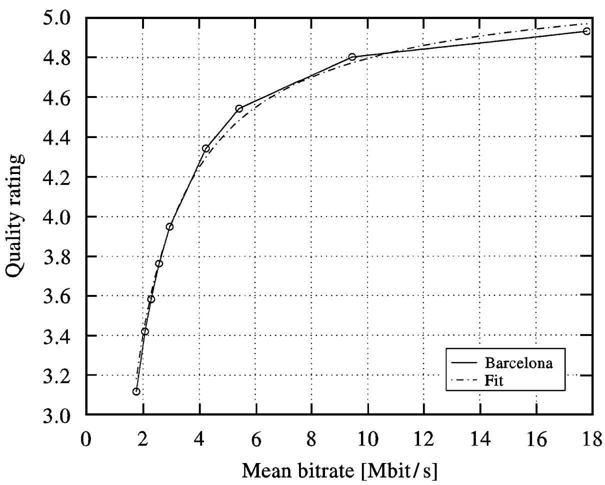

Figure 9. MPQM video quality versus average output encoding bit rate for the Barcelona sequence. Fitting parameters for Eqn (3): ($\chi_Q = -0.045$, $Q_0 = 5.225$, $\chi_R = 124.761$, $\xi_R = 1.116$).

Data loss impact on video quality

Up to this point, we did not consider any data loss in the video stream. Figure 10 illustrates how the video quality is affected by uniformly distributing TS packet losses over a 400-frame long MPEG-2 transport stream[6]. It is shown that, on a semi-logarithmic scale and for a given *MQUANT* (average bit rate), first the video quality remains constant with the *PLR*. This constant value corresponds to the encoding quality (Q_E). Then, beyond a certain *PLR*, the perceptual quality quickly drops.

The higher the average bit rate, the lower the PLR after which the video quality drops, and inversely. The *PLR* is indeed defined as the number of lost packets per time unit divided by the number of packets transmitted during the same time unit. In MPEG-2 video delivery, the packet size does not depend on the encoding bit rate [17, 18]. Therefore, the higher the encoding bit rate, the higher the number of packets transmitted per time unit.

[6] It turned out to be meaningless to perform data loss simulations over a smaller set of frames.

Thus, for a given PLR, the higher the encoding bit rate, the higher the number of packets lost per time unit[7].

Hence, the relation between video quality Q and PLR may be represented by a straight line on a linear scale:

$$Q = QE + \chi_L^* PLR, \tag{4}$$

where Q_E corresponds to the encoding quality (given by Eqn (3)) and χ_L^* depends on both the complexity of the sequence and the average bit rate. In other words, for a given sequence and a fixed $MQUANT$, the video quality, averaged over the whole sequence, linearly decreases with the PLR.

This relation still holds if we multiply the PLR by a constant. We observed that, for a given $MQUANT$, the relation between the end-to-end video quality Q and the product $\bar{R} \times PLR$ could be well approximated by a straight line of slope χ_L. Therefore, Eqn (4) becomes:

$$Q = Q_E + \chi_L(\bar{R} \times PLR), \tag{5}$$

where χ_L is almost independent of the $MQUANT$ especially for low to medium bit rates.

It is to be noted that PLR after which the quality quickly drops may be shifted to the right by means, for instance, of error correction mechanisms (e.g. syntactic protection [19], FEC [20]).

Joint Impact of MPEG-2 Rate and Data Loss on Quality

In this section, we demonstrate why a joint analysis of the impact of both the MPEG-2 encoding bit rate and the data loss ratio on the video quality is the only way to get correct conclusions. We explain, for example, why the video quality may decrease when the encoding bit rate is increased in an error-prone environment.

Joint impact analysis

As stated at the end of the previous section, the PLR and the encoding bit rate (packet rate) are intimately related to each other in regards to their impact on video quality. For example, the higher the bit rate, the higher the encoding quality (up to saturation) but the lower the PLR after which the video quality quickly drops, and conversely. Therefore, the relation between quality and encoding bit rate for a non-zero PLR should somehow exhibit an optimum value. This behavior is illustrated in Figure 11. We indeed see that the video quality first increases (encoding quality) with the average bit rate and then decreases after around 4 MBits/s for the "Barcelona sequence" (data loss). This optimal average bit rate directly depends on the video scene type. We observed that it was fairly independent of the PLR though. Furthermore, we can extend the bit rate range over which the quality is optimal by implementing mechanisms such as adaptive quantization and/or FEC-based protection, as mentioned herebefore.

Such a result is crucial for the design of network-aware rate controllers or efficient error concealment algorithms. The former algorithm would consist in adjusting the encoding rate ($MQUANT$) in order to follow the optimal working point in a network environment where the PLR may significantly vary over a video transmission duration (e.g. broadcasting via radio links, the Internet network).

Tri-dimensional representation

The purpose of this subsection is to put all results together and represent them by a single graph. Thus, by putting together Eqn (3) and Eqn (5), we obtain a good model of the end-to-end video quality Q:

$$Q = 5 + \chi_Q \times \left(\frac{\bar{R}}{\chi_R}\right)^{-\frac{1}{\xi_R}} + \chi_L + (\bar{R} \times PLR), \tag{6}$$

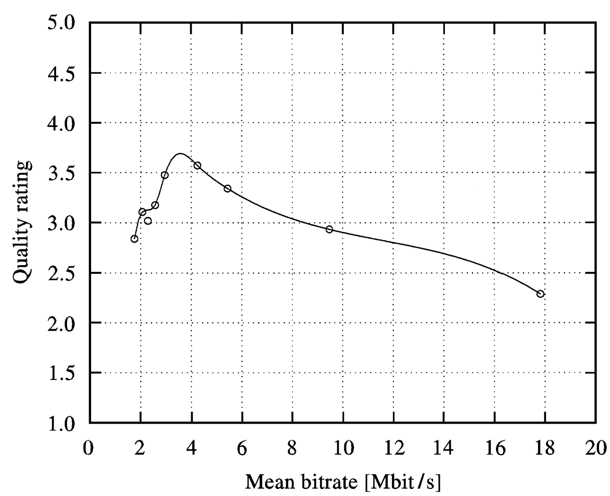

Figure 11. $MPQM$ versus average encoding bit rate for $PLR = 5 \times 10^{-3}$ for the Barcelona sequence.

[7] The number of video frames transmitted per time unit is independent of the encoding bit rate.

in which the two first terms of the sum represent the encoding quality (see Eqn (3)).

We then performed a complete set of measurements in order to verify this relation. The same simulation setup as the one presented in the previous section has been used. Figure 12 presents the resulting surface for the "Barcelona" sequence while Figure 13 shows its corresponding fit using Eqn (6).

Several results may be extracted from these graphs. Most of these results have already been discussed throughout this paper. In general, when considering video transmission over lossy networks, not only it is bandwidth consuming to increase the encoding bit rate above a certain threshold due to saturation of quality (which varies according to the scene complexity), it may also be quality consuming. In other words, when the user-oriented QoS is not high enough, an increase of the encoding bit rate at a fixed PLR may even degrade the quality, depending on the position of the working point on the 3D graph presented herebefore. These is an optimal bit rate to be determined that maximizes the end-user perception of the service under certain given network conditions (i.e. network impairments).

Conclusion and Future Works

The combined effect of the coding bitrate and the network impairments on the user-perceived quality is still not well understood. However these results are needed for the design and deployment of packet video services. One of the common misleading intuitions is that increasing the coder bit rate enhances image quality. In this paper we have shown that this intuition is proper to a lossless communication channel and that the quality-rate function is no longer a strictly increasing function when video packets are subject to loss.

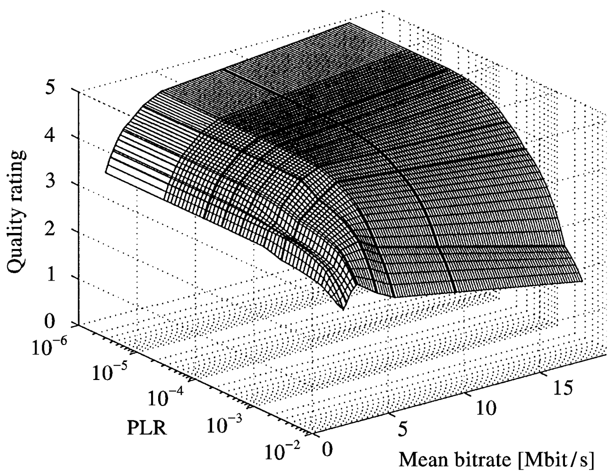

Figure 12. Q versus average bit rate and PLR: Simulations on the Barcelona sequence.

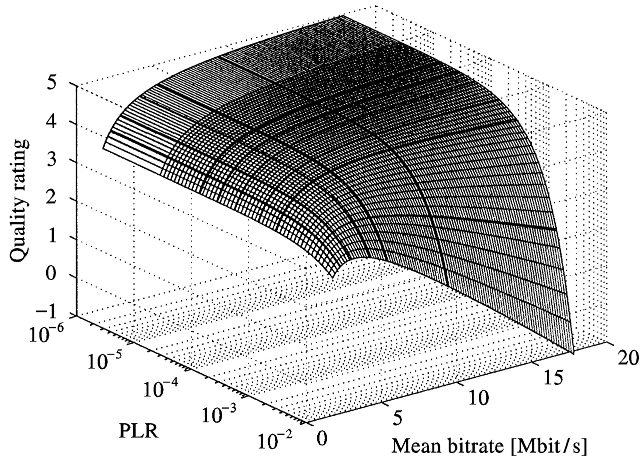

Figure 13. Q versus average bit rate and PLR: Fitting function with parameters: ($\chi_Q = -0.045$, $Q_0 = 5.22$, $\chi_R = 124.76$, $\xi_R = 1.12$ and $\chi_L = -33.9$).

The major conclusion is that image quality cannot be improved by acting on the coding bit rate only: increasing the bit rate above a certain threshold results in quality degradations. For a given packet loss ratio, there is a quality-optimal coding rate that has to be found. Although the relationship between coding bit rate, packet loss ratio and user-level quality is intrinsically complex, it can be characterized by a simple expression and parameters set. These parameters seem to depend on properties of the video scene (e.g. encoding complexities). They have to be predicted when video is coded and transmitted in real-time over lossy networks. Therefore, this work is being extended to on-line prediction of the 3D quality graph in the context of MPEG-2, as well as other emerging encoding standards.

References

1. Karlsson G. (1996) Asynchronous Transfer of Video. *IEEE Communications Magazine*, **34**: 118–126.
2. Verscheure O., Garcia X., Karlsson G. & Hubaux, J.-P. (1998) User-Oriented QoS in Packet Video Delivery. *IEEE Network Magazine*, Nov. Dec. 1998.

3. ISO/IEC JTC 1 (1996) *Information Technology — Generic Coding of Moving Pictures and Associated Audio Information — Part 1, 2 and 3*, ISO/IEC JTC 1.
4. Haskell, B.G., Puri A., & Netravali, A.N. (1997) *Digital Video: an Introduction to MPEG-2*. Digital Multimedia Standards Series. London Chapman and Hall.
5. Webster, A., Jones, C., Pinson, M., Voran, S. & Wolf, S. (1993) An Objective Video Quality Assessment System Based on Human Perception. In: *SPIE — Human Vision, Visual Processing and Digital Display*, Vol. 1913, pp. 15–26.
6. Lubin, J. & Fibush, D. (1997) *Sarnoff JND Vision Model*, T1A1.5/97-612, Contribution to T1 Standards project, August 1997.
7. van den Branden Lambrecht, C. J. (1996) *Perceptual Models and Architectures for Video Coding Applications*, Ph.D. Thesis, Signal Processing Laboratory, Swiss Federal Institute of Technology, CH 1015 Lausanne, Switzerland.
8. van den Branden Lambercht C.J. & Verscheure, O. (1996) Perceptual Quality Measure using a Spatio-Temporal Model of the Human Visual System. In: *Proceedings of the SPIE,* San Jose, CA, January 28–February 2 1996, Vol. 2668, pp. 450–461.
9. van den Branden Lambrecht, C.J. (1996) A Working Spatio-Temporal Model of the Human Visual System for Image Restoration and Quality Assessment Applications. In: *ICASSP*, Atlanta, GA, May 1996, pp. 2293–2296.
10. Ardito, M., Barbero, M., Stroppiana, M., Visca M. (1994) Compression and Quality. In: *Proceecings of the International Workshop on HDTV 94,* Brisbane, Australia, October 1994.
11. Fogg, C. (1996) mpeg2encode/mpeg2decode. In: *MPEG Software Simulation Group.*
12. Roberts, J.W., Guibert, J. & Simonian, A. (1991) Network Performance Considerations in the Design of a VBR Codec. In: *Queuing Performance and Control in ATM* (Cohen, J.W. & Pack, C.P. eds) pp.77–82.
13. Verscheure, O. & van den Branden Lambrecht C.J. (1997) Adaptive Quantization Using a Perceptual Visibility Predictor. In: *IEEE ICIP.*
14. Sakazawa, S., Yakishima, Y., Wada, M & Hatori, Y. (1996) Coding Control Scheme for a Multi-Encoder System. In: *7th International Workshop on Packet Video,* March 1996, pp. 83–88.
15. Fukuda, K., Wakamiya, N., Murata, M. & Miyahara, H. (1998) On Flow Aggregation for Multicast Video Transport. In: *6th IFIP International Workshop on Quality of Service (IWQoS),* May 1998.
16. Verscheure, O., Basso, A., El-Maliki, M. & Hubaux, J.-P. (1996) Perceptual Bit Allocation for MPEG-2 Video Coding. In: *ICIP,* Lausanne, Switzerland, September 16–19, 1996.
17. ITU-T, Ed. (1998) *Recommendation H.222.1 Multimedia Multiplex and Synchronization for Audiovisual Communications in ATM Environments,* ITU-T, Geneva, March 1998.
18. Goyal, V., Hoffman, D., Fernando, G., Civanlar, R. (1998) RTP Payload Format for MPEG1/MPEG2 Video. Tech. Rep. RC 2250, IETF, January 1998.
19. Frossard, P. & Verscheure, O. (1998) AMIS: Adaptive MPEG-2 Information Structring. In: *SPIE International Symposium on Voice, Video, and Data Communications,* Boston, USA, November 1998.
20. McAuley, A.J. (1990) Reliable broadband communications using a burst erasure correcting code. In: *ACM SIGCOMM'90*, Philadelphia, USA.