Objective Evaluation of the Perceived Quality of Video Content

H. Koumaras, G. Gardikis, A. Kourtis

Institute of Informatics and Telecommunications NCSR «DEMOKRITOS»,
Patriarchou Gregoriou Str., Agia Paraskevi, Attiki, 15310 Athens Greece.
National Technical University of Athens, Mobile Radiocommunications Lab,
9 Iroon Polytechneiou Str., Zografou, 15773 Athens Greece.
Email: koumaras@iit.demokritos.gr, gardikis@iit.demokritos.gr, kourtis@iit.demokritos.gr

ABSTRACT

This paper presents a method for the objective evaluation of the perceived quality of video content. The proposed method is based on the experimental measurement of a simple metric derived from the encoding and playback of a short video clip in MPEG-4 format. Implementation and tests have been performed in the frame of the IST/SOQUET project

Keywords: Perceived Quality of Service (PQoS), MPEG-4 Objective measurements

I. INTRODUCTION

The mobile communications systems of the 3rd generation, among which UMTS is the most dominant, will provide the platforms for a multitude of novel services and applications. Where mobile communications systems of the 2nd generation concentrated on voice and data services, the 3rd generation systems will provide a broader range of services. Multimedia, including Internet content, downloading of short video clips etc. is expected to be a significant part of the new offerings. One of the 3G visions is that services will be sold in a consumer mass market based on the provision of content at a requested quality. There are a number of approaches to this, one being the use of Perceived Quality of Service (PQoS) concept. The evaluation of the PQoS for multimedia and audiovisual content that have variable bandwidth demands, will provide a user with a range of choices covering for example the possibilities of low, medium or high quality connections, indication of service availability and costs. Based on the available radio resources a quality map can be computed and will be provided as a menu of options by service providers to end users grouped into different connection categories or some other rational partitioning. Dissemination of the menu of services is the responsibility of the service providers (e.g. via a Web site). End users choose the QoS that they want to start the session with, so that the service provider can launch the service with a set of parameters that match those requested by the user. The application of the PQoS concept gives the operator (or even the service provider) the capability of a better exploitation of the network resources, because it is possible to allocate only those radio resources sufficient to maintain user satisfaction. Since this is true for each user, then the overall allocation of the radio resource results in an optimization of the system spectral efficiency.

The Information Society Technologies (IST) project SOQUET (IST-2000-28521) aims to promote the adoption and usage of the perceived quality of service concept in upcoming 3G networks. In this context, one of its major objectives is to develop efficient and cost-effective solutions for the delivery of the envisaged services at a requested (user perceived) quality. This is essential for those services targeting a consumer mass market (existing and new QoS-based interactive services) and will in turn justify the investment to such services and technologies for the business entities involved in a 3G (and beyond) venture.

In order to provide audiovisual content at different perceived quality levels, it is necessary to define the variation of PQoS as an equation of the encoding bit rate. Precise curves of PQoS vs. bit rate can be derived using an audience of people, who are watching the video (e.g. a short video clip) and score its quality, as perceived by them. However, this procedure is expensive and time consuming, because it involves a lot of people, gathering and elaboration on results to derive mean values. For actual commercial applications, objective measurement of PQoS for any type of video content is required. In this respect, this paper defines a simple metric for the measurement of the PQoS level and presents a method that is capable of making objective measurements of the perceived quality of audiovisual content, in the form of a short video clip.

II. DEFINITION OF METRICS FOR EVALUATING PQoS

Among the standardized digital video encoding formats (MPEG-1/2/4), the MPEG-4 is mostly preferred in the distribution of interactive multimedia services over IP, while MPEG-2 is almost exclusively used in DVB networks. Furthermore, MPEG-4 is most suitable for 3G networks, because it provides a good quality at a lower bit rate, compared to the other two formats given that UMTS can reach up to 2 Mbit/sec maximum bit rate [1],[5]. Unlike MPEG-1 and MPEG-2, the evolution of MPEG-4 has followed two paths; the former is based on Windows Media Technologies (WMT) and the latter on ISO/IEC N4668 standard (March 2002).

The most reliable method to measure the perceived quality level of audiovisual content is to use subjective evaluations. Curves of PQoS vs. bit rate can be derived using an audience of people, who are watching the video (e.g. a short video clip) and score its quality, as perceived by them. Such curves are shown in figure 1 [2]. Curve (A) represents a video clip with low temporal and spatial dynamics, i.e. whose content has “poor” movements and low picture complexity. Such a curve can be derived, for example from a talk show. Curve (C) represents a short video clip with high dynamics, such as a football match. Curve (B) represents an intermediate case. Each curve -and therefore each video clip- can be characterized by: (a) the low bit rate (BR_L), which corresponds to the lower value of the accepted PQoS (PQ_L) by the audience, (b) the high bit rate (BR_H), which corresponds to the minimum value of the bit rate for which the PQoS reaches its maximum value (PQ_H) (see BR_H for curve (A) in figure 1) and (c) the mean inclination of the curve, which can be defined as $ME=(PQ_H - PQ_L) / (BR_H - BR_L)$. From the curves of figure 1, it can be deduced that video clips with low dynamics have lower BR_L and higher ME than clips with high dynamics.

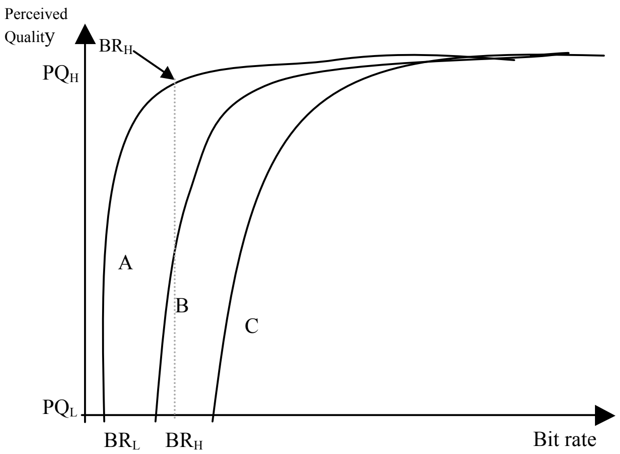

Figure 1. Results of subjective evaluation tests

A general analytical expression for the curves in figure 1 has not yet been derived and there is still work to be done in theoretical level. Furthermore, the fact that these curves are derived from subjective tests, deprives their use in practical and commercial systems. A simplified approach to these curves in conjunction with a practical technique to evaluate them objectively, would be very useful for applications in commercial systems.

The method proposed in this paper is based on a linear approximation of the curves of figure 1. Such an approach is depicted in figure 2. Although it is evident that there is a deviation between the actual and linearly approximated curves, this error is not significant in comparison with the practical benefit of the proposed method.

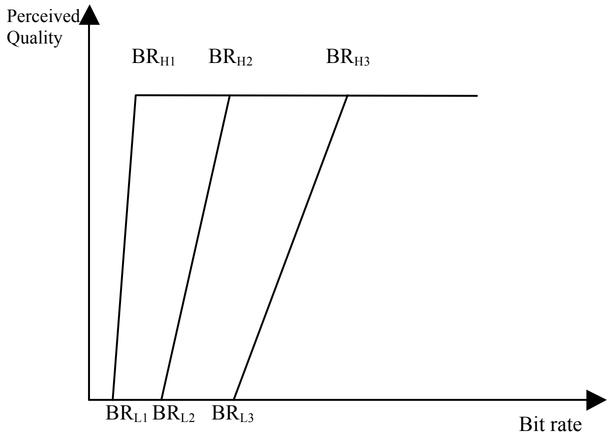

Figure 2. The linear approximation of subjective evaluations

In the proposed approach, two metrics are required to be measured: the BR_L and the ME of the corresponding curve. The measurement of these two metrics is enough to define the bit rate that corresponds to any value of the perceived quality for a video clip. The following section presents a technique to estimate these two metrics objectively.

III. OBJECTIVE ESTIMATION OF BR_L AND ME

For the needs of this paper, five test video clips with different content dynamics each were used. These well known video clips have been developed by Tektronix Corp. [7] and Sanoff Laboratories and are shown on Table 1. Each test video clip was transcoded from its original MPEG-2 format at 12 Mb/s to WMT MPEG-4 format, at different bit rates. For each bit rate, a different MPEG-4 file was created. CIF resolution (352x288) and 25 frames per second (fps) were common parameters for the transcoding process in all test videos.

Each MPEG-4 video clip was then played back using a suitable media player, featuring an option to measure the mean value of fps over the whole duration of the video clip (statistics option).

During playback, it was observed that the perceived quality decreases as the bit rate drops, even though the mean frame rate remain constant to 25 fps, as measured by the player statistics tool. It was also observed that when the bit rate drops below a certain threshold, picture pauses appear. These can be measured by the mean frame rate, which drops below 25 fps, as measured by the statistics tool.

The variation of the mean frame rate during playback, versus encoding bit rate is shown in figure 3, for the various test video clips. Note that since the statistics tool measures mean values, the points on the curves do not correspond always on integer values.

Clip 1	Suzie
Clip 2	Cactus & Comb
Clip 3	Table Tennis
Clip 4	Flowers & Mile
Clip 5	Mobile & Calendar

Table 1. The five test video clips

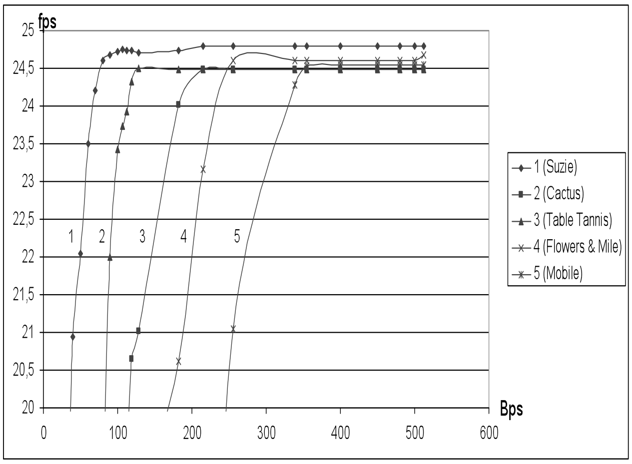

Figure 3. The mean fps vs. bit rate curves

From these curves it is evident that the value of the bit rate, below which the mean frame rate drops below its original value (25 fps), depends on the dynamics of the video content of the clip. In other words, video clips with high dynamics (like mobile & calendar) require higher bit rate to keep the mean fps value at 25, than video clips with lower dynamics (like Suzie). Moreover, the mean inclination of the curves of figure 3 is lower for clips with high dynamics, exactly as in the curves of figure 1 and figure 2.

In order to correlate curves of figures 2 and 3, the following assumptions can be made:

(a) When the frame rate drops below 25 fps, the perceived quality is considered to be unacceptable. So, it is possible to consider the BR_L point of each curve in figure 2 is equal to the objectively measured bit rate in figure 3, where the curve starts to bend. Both values correspond to the minimum level of accepted perceived quality.

(b) The inclination of each curve of figur*e* 3 is equal to the ME of the corresponding curve of figure 2.

The above assumptions allow to use the objectively measured metric, estimated from the curves of figure 3, to measure the PQoS level for any type of video content.

Observing the curves of figure 3, it can be deduced that there is a correlation between BR_L and the mean inclination of each curve. Table 2 shows the mean inclination values vs. BR_L for the video clips of Table 1.

Interpolating the values of table 2 it is possible to use just a single objectively measured metric (the bit rate for which the mean frame rate drops below 25 fps) for any short video clip, in order to estimate the variation of PQoS versus bit rate, for that video clip.

Video Dynamics	BRL (Kbps)	Mean Inclination (fps/kbps)
Low	80	0.091
	130	0.083
Medium	210	0.056
	250	0.050
High	400	0.030

Table 2. The Mean Inclination vs. BR_L bit rate

IV. IMPLEMENTATION OF THE PROPOSED METHOD.

The main scope of SOQUET project is to develop efficient and cost-effective solutions for the delivery of the envisaged services at a requested (user perceived) quality. This means that the service provider offers a number of different qualities, among which the consumer is able to choose the one, at which he/she prefers to watch the offered video clips. In order to achieve this, the service provider has to define the curve of PQoS vs. bit rate for each video clip it intends to offer. Then it has to encode each video clip with the encoding bit rate that corresponds to every level of PQoS.

According to the above model, and assuming 5 quality levels (1=lowest and 5=highest), it is possible to define the values of bit rate that correspond to quality levels 1 to 5 and given a short video clip, it is sufficient to define the bit rate for which the mean frame rate drops below 25 fps for this clip. This corresponds to the lowest accepted perceived quality level (level 1). Then, using the data of table 2 it is possible to estimate the encoding bit rates, which correspond to the quality levels (e.g. 5 levels). Finally, the service provider stores the encoded clips with different quality levels into a server, from where end users are able to access them.

In order to define the bit rate that corresponds to quality level 1 (lowest accepted perceived quality) for a given short video clip it is necessary to encode the clip at a significant number of bit rates, according to those shown in figure 4. This is a time consuming process. A simplified approach is to divide the axis of bit rate in N equal areas, as shown in figure 4.

Each area is defined by its initial value, e.g. area 3 is defined by BR_{L3}. Given a short video clip, the process starts with its encoding in MPEG-4 at n different bit rates: BR_{L1}, BR_{L2}, … BR_{LN} according to the areas shown in figure 4. Then, the mean fps for each bit rate is measured using the statistics tool. This allows the determination of the bit rate for which frame loss begins (quality level 1). From table 2, it is possible to define the values of the bit rates that correspond to various quality levels.

Observing curves of figure 4 that correspond to BR_{L1} and BR_{Lx} it is derived that video clips with high dynamics can be offered with more quality level choices than videos with low dynamics.

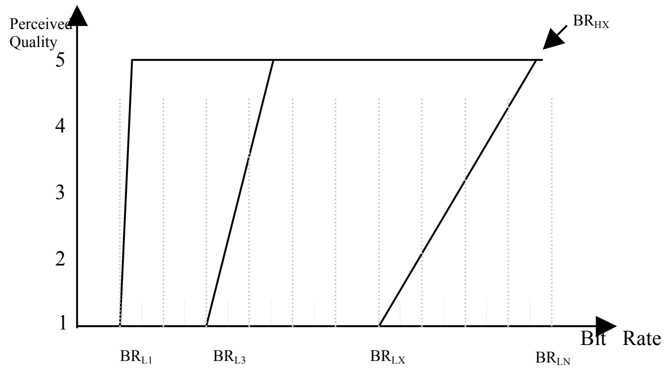

Figure 4. The proposed method in *SOQUET* project

For example a video clip with low temporal and spatial dynamics is useless to be offered in many quality levels, because the differences will be negligible. One or two quality levels are enough for a static video, i.e. a talk show. In other words, if the distance between the BR_{Lx} and BR_{Hx} (see figure 4) of a clip is large (this indicates video clips with high dynamics) then the offering quality levels will be many. For video clips with low dynamics this distance will be short and the offering quality levels few.

When a user is about to choose the quality level, at which he/she wishes to watch a specific video clip, then the provider should offer as possible choices all that levels, which are implemented by pre-selected bit rates that are contained in the area [BR_{Lx}, BR_{Hx}], of the selected video clip. Of course the provider will define the exact number of these bit rates, according to the above note that video clips with high dynamics are expected to be offered with more quality level choices than videos with low dynamics. Then using the proposed method, the provider can precisely find the bit rates that correspond to specific quality levels.

V. CONCLUSIONS

The provision of audiovisual content at different perceived quality levels requires curves of PQoS vs. bit rate, which are usually derived via an audience of people. This paper defines a simple metric for the quantitative estimation of the PQoS level and presents a method that is capable of making objective measurements of the perceived quality of audiovisual content, in the form of a short video clip. The proposed method is very useful for actual commercial applications and it has been implemented and tested in the frame of IST/SOQUET project.

ACKNOWLEDGEMENTS

The work in this paper was carried out in the frame of the Information Society Technologies (IST) project SOQUET (IST-2000-28521).

REFERENCES

[1] ISO-IEC/JTC1/SC29/WG11 N2604 "Report Of the Formal Verification Tests on MPEG-4 Video Error Resilience"

[2] Z. He, S. K. Mitra, "Optimum Bit Allocation and Accurate Rate Control for Video Coding via p-Domain Source Modeling" IEEE *Transactions on Circuits and Systems for Video Technology VOL.12 NO.10 October 2002*

[3] F. Pereira, "MPEG-4 Video Subjective Test Procedures and Results" " IEEE *Transactions on Circuits and Systems for Video Technology VOL.7 NO.1 February 1997*

[4] ISO-IEC/JTC1/SC29/WG11 N4668 "MPEG-4 Overview" March 2002

[5] Z. He, Y. Kim, S. K. Mitra, "A novel linear source model and a unified rate control algorithm for H.263/MPEG-2/MPEG-4", presented at the Int. Conf. Acoustics, Speech and Signal Processing, Salt Lake City, UT, May 2001

[6] ISO-IEC/JTC1/SC29/WG11 W2711 "Report of the Formal Verification Tests on MPEG-4 Coded Based Coding", March 1999

[7] Tektronix Web Site, http://www.tektronix.com

[8] D. LeGall, "MPEG: A video compression standard for multimedia application", *Commun. ACM*, vol. 34,pp.46-58, Apr.1991

[9] "Final report from the video quality experts group on the validation of objective models of video quality assessment" March 2000 report

[10] St. Wolf, M. H. Pinson, "Spatial – Temporal Distortion Metrics for In-service Quality Monitoring of Any Digital Video System", Institute for Telecommunication, National Telecommunications and Information Administration