End-to-End Prediction Model of Video Quality and Decodable Frame Rate for MPEG Broadcasting Services

Harilaos Koumaras, Anastasios Kourtis
Institute of Informatics and Telecommunications
NCSR Demokritos
Athens, Greece
{koumaras, kourtis}@iit.demokritos.gr

Cheng-Han Lin, Ce-Kuen Shieh
High Performance Parallel & Distributed Syst. Laboratory
National Cheng Kung University
Tainan, Taiwan
{jhlin5,shieh}@hpds.ee.ncku.edu.tw

***Abstract*—This paper proposes, describes and evaluates a novel theoretical framework for end-to-end video quality assessment of MPEG-based video services in hand-held and mobile wireless broadcast systems. The proposed framework consists two discrete models: A model for predicting the video quality of an encoded signal at a pre-encoding state by specifying the bit rate that satisfies a specific level of user satisfaction and a model that maps the packet loss ratio of the transmission channel to the quality degradation percentage of the broadcasting service. The accuracy of the proposed framework is experimentally validated, demonstrating its efficiency.**

***Keywords*— Video Quality, MPEG, Packet Loss, DVB.**

I. Introduction

RECENTLY, MPEG-based applications that are specialized and adapted in broadcasting digitally encoded audiovisual content have known an explosive growth in terms of development, deployment and provision. The new era of digital video broadcasting for hand-held terminals has arrived and the beyond MPEG-2 based transmission for terrestrial or satellite receivers is a fact, setting new research challenges for the assessment of Perceived Quality of Service (PQoS) under the latest MPEG-4/H.264 and the DVB-H standard.

MPEG standards exploit in their compression algorithms the high similarity of the depicted data in the spatial, temporal and frequency domain among subsequent frames of a video sequence. Removing the redundancy in these three domains, it is achieved data compression against a certain amount of visual data loss, which from the one hand it cannot be retrieved but on the other hand it is not perceived and conceived by the mechanisms of the Human Visual System.

Therefore, MPEG-based coding standards are characterized as lossy techniques, since they provide efficient video compression with cost a partial loss of the data and subsequently the video quality degradation of the initial signal. Due to the fact that the parameters with strong influence on the video quality are normally those, set at the encoder (with most important the bit rate), the issue of the user satisfaction in correlation with the encoding parameters has been raised.

A content/service provider, depending on the content dynamics, must decide for the configuration of the appropriate encoding parameters that satisfy a specific level of user satisfaction.

Currently, the determination of the encoding parameters that satisfy a specific level of video quality is a matter of recurring subjective or objective video quality assessments, each time taking place after the encoding process (repetitive post-encoding evaluations). Subjective quality evaluation processes of video streams require large amount of human resources, establishing it as an impractical procedure for a service provider. Similarly, the repetitive use of objective metrics on already encoded sequences may require numerous test encodings for identifying the specified encoding parameters, which is also time consuming and financially unaffordable from a business perspective.

Once the broadcaster has encoded appropriately the offered content at the preferred quality level, then the provision of the service follows. Digitally video encoded services, due to their interdependent nature, are highly sensitive to transmission errors (e.g. packet loss, network delay) and require high transmission reliability in order to maintain between sender and receiver devices their stream synchronization and initial quality level. Especially, in video broadcasting, which is performed over wireless environments, each transmitted from one end video packet can be received at the other end, either correctly or with errors or get totally lost. In the last two cases, the perceptual outcome is similar, since the decoder at the end-user usually discards the packet with errors, causing visual artifact on the decoded frame and therefore quality degradation.

In this context, the paper aims at proposing, describing and evaluating a theoretical framework for end-to-end video quality assessment of MPEG-based broadcasting services, focusing on:

i. The prediction of the encoding parameters that satisfy a specific video quality level in terms of encoding bit rate and content dynamics.
ii. The mapping of the packet loss ratio during the transmission to the respective quality degradation percentage.

Through the proposed end-to-end video quality assessment framework, the content provider (i.e. the broadcaster) will be able to estimate the finally delivered video quality level, considering specific encoding parameters and transmission conditions. Such an end-to-end perceived QoS framework will not only play an essential role in performance analysis, control and optimization of broadcasting systems, but it will also contribute towards a more efficient resource allocation, utilization and management.

The rest of the paper is organized as follows: Section II performs a literature review on the relative research works, focusing both on the video quality assessment and the estimation of the degradation due to the conditions of the transmission channel. Also, in this section are described the fundamental concepts of a MPEG encoded signal, which will be later used for the description of the proposed framework. Section III describes and evaluates the proposed model for the prediction and determination of the encoding bit rate value that satisfies a specific level of user satisfaction. Similarly, Section IV discusses the consequence of a packet loss on the transmitted broadcasting signal, focusing on the decoding performance of the service. In this context, it is described the proposed model for the video quality degradation over error-prone transmission channel. In section V, the concept of the end-to-end video quality assessment framework is introduced, described and explained. Finally, Section VI concludes the paper.

II. BACKGROUND

A. *Video Quality Assessment Methods*

The advent in the field of video quality assessment is the application of pure error-sensitive functions between the encoded and the original/uncompressed video sequence. These primitive methods, although they initially provided a quantitative approach of the degradation caused by the encoding procedure, practically they do not reflect reliably the video quality as it is observed and perceived by human viewers.

Beyond these primitive models, currently the evaluation of the video quality is a matter of objective and subjective procedures, which are applied after the encoding process (post-encoding evaluation).

The subjective test methods, which have mainly been proposed by International Telecommunications Union (ITU) and Video Quality Experts Group (VQEG), involve an audience, who watch a video sequence and score its quality as perceived by the participants, under specific and controlled watching conditions. Afterwards, usually the Mean Opinion Score (MOS) is exploited for the statistical analysis and processing of the collected data.

Subjective video quality evaluation processes require large amount of human resources, making it a time-consuming process (e.g. large audiences evaluating test sequences). On the other hand, objective evaluation methods provide faster quality assessment, exploiting multiple metrics related to the encoding artifacts (e.g. tilling, blurriness, error blocks, etc).

The majority of the objective methods require the undistorted video source as a reference entity in the quality evaluation process. Due to this requirement, they are characterized as Full Reference (FR) Methods [1-3]. However, it has been reported that these complicated FR methods do not provide more accurate results than the simple mathematical measures (such as PSNR). Towards this, lately some novel full reference metrics have been proposed based on the video structural distortion and content entropy [5-8].

On the other hand, the fact that these methods require the original video signal as reference deprives their use in broadcasting services, where the initial undistorted clips are not accessible at user side. Moreover, even if the reference clip becomes somehow available, then synchronization predicaments between the undistorted and the distorted signal (which may have experienced frame losses) make the FR methods practically inapplicable.

Due to these reasons, the recent research has been focused on developing methods that can evaluate the PQoS level based on metrics, which use only some extracted structural features from the original signal (Reduced Reference Methods) [9-13] or do not require any reference video signal (No Reference Methods). The NR methods can be classified into two classes: The NR-visual based and the NR-coded based. The first methods must initially decode the bit stream and estimate the video quality at the visual domain [14-19], while the second ones assess the perceived quality directly through the compressed bit stream, without requiring any decoding [20-25].

Finally, some alternative objective methods have been proposed, which move beyond the simple post-encoding quality assessment and introduce the concept of video quality prediction for given encoding parameters and content dynamics at a pre-encoding state [26-28, 40]. In this direction will focus the content of this paper and more specifically the proposed model for the determination of the bit rate values that satisfy specific perceptual levels.

B. *Quality Degradation due to Transmission Errors*

The issue of mapping the perceptual impact of transmission errors (like packet loss) during the broadcasting on the delivered perceptual video quality at the end-user is a fresh topic in the field of video quality assessment since the relative literature appears to be limited with a small number of relative published works.

In this framework, S. Kanumuri ***et al*** [29] proposed a very analytical statistical model of the packet-loss visual impact on the decoding video quality of MPEG-2 video sequences, specifying the various factors that affect the perceived video quality and visibility (e.g. Maximum number of frames affected by the packet loss, on what frame type the packet loss occurs etc). However, this study focuses mainly on the pure study of the MPEG-2 decoding capabilities, without considering the parameters of the digital broadcasting or the latest encoding standards.

Similarly, in [30] is presented a transmission distortion model for real-time video streaming over error-prone wireless networks. In this work, an end-to-end video distortion study is performed, based on the modeling of the impulse propagation error (i.e. the visual fading behavior of the decoding artifact). The deduced model, although it is very accurate and robust, enabling the media service provider to predict the transmission distortion at the receiver side, is not a generic one. On the contrary, it is highly dependent on the video content dynamics and the selected encoder settings. More specifically, it is required an initial quantification of the spatial and temporal dynamics of the content, which will allow the appropriate calibration of the model. This prerequisite procedure (i.e. adapting the impulse transmission distortion curve based on the least mean square error criteria) is practically inapplicable by an actual content creator/provider. Moreover, the strong dependence of the proposed model on the spatiotemporal dynamics of the content deprives its implementation on sequences with long duration and mixed video dynamics, since not a unique impulse transmission distortion will be accurate for the whole video duration.

In this context, our paper describes, proposes and testes a generic model for end-to-end video quality prediction for MPEG-based broadcasting services. Our framework consists two discrete parts:

- A method for predicting and specifying for a given content the encoding parameters that satisfy a specific perceptual level at a pre-encoding state
- A model of the perceptual impact of the broadcasting packet loss ratio on the delivered perceived quality of the transmitted service.

Thus, to the best of our knowledge, this work is one of the first models providing end-to-end video quality prediction across all the lifecycle of the media content: From the service generation down to the content consumption at the viewer side.

C. *MPEG Video Structure*

The MPEG standard [31] defines three frame types for the compressed video streams, namely I, P and B frames. The I frames are also called Intra frames, while B and P are known as Inter frames. The successive frames between two succeeding I frames is defined as Group Of Pictures (GOP). The frame classification is mainly based on the procedure, according to which each frame type has been generated and encoded. This differentiation sets also some special requirements for the successful decoding of each frame type.

More specifically, MPEG I frames (Intra-coded frames) are encoded independently and there is no special requirement in their decoding process, given that all the respective data packets have been successfully received. The encoding of the MPEG P frames (Predictive-coded frames) is based on reference spatial areas from the preceding I or P frames within the specific GOP. Therefore, for their successful decoding -except for the successful reception of their respective data- it is required successful decoding of the reference I or P frames. Finally, MPEG B frames (Bi-directionally predictive-coded frames) are encoded using references from the preceding and succeeding I or P frames. Consequently, for their successful decoding apart from the successful reception of the data packets that carry the B frame, also the respective reference frames must be successfully received and decoded.

The structure of the GOP is generally specified by the selected encoding settings. In the MPEG literature the GOP pattern is described by two parameters GOP(N,M), where N defines the GOP length (i.e. the total number of frames within each GOP) and the M-1 is the number of B frames between I-P or P-P frames. For example, as shown in figure 1, GOP(12, 3) means that the GOP consists one I frame, three P frames, and eight B frames. Also seen in figure 1, the second I frame marks the beginning of the next GOP. The arrows indicate that the B and P frames successful decoding depends on the respective preceding and succeeding I or P frames.

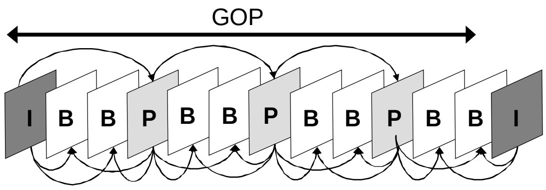

Fig. 1. A sample of MPEG GOP (N=12, M=3)

Therefore, from the hierarchical structure of MPEG encoding as it is depicted on figure 1, a video frame may be considered as directly or indirectly undecodable. Direct undecodable is considered a video frame when there are not enough received packets of the frame in order to achieve a successful decoding. On the other hand, indirect undecodable is considered a video frame when a reference frame is directly or indirectly undecodable. For simplicity, we do not consider video concealment issues in this study and we set the Decodable Threshold (DT) [13] equal to 1.0. Therefore, our analysis provides the worst-case scenario in terms of video quality degradation and decoding robustness.

III. Modeling and Predicting Video Quality

In digital video encoding the Block Discrete Cosine Transformation (BDCT) is exploited, since it exhibits very good energy compaction and de-correlation properties. In this paper, we use the following conventions for video sequences: Every real NxN frame f is treated as a N^2x1 vector in the space R^{N^2} by lexicographic ordering either in rows or columns.

The DCT is considered as a linear transform from $R^{N^2} \to R^{N^2}$. Thus, for a typical frame f, we can write:

$$F = Bf$$

Since B matrix is unitary, the inverse DCT can be expressed by B^t, where t denotes the transpose of a vector or matrix. Thus, the inverse transform can be described as:

$$f = B^t F$$

The elements of frame $F = Bf$ in the frequency domain can be expressed as the coefficients of the vector f, using the DCT basis in R^{N^2}. Thus

$$f = \sum_{n=1}^{N^2} F_n e_n$$

where e_n is the normalized DCT basis vector and F_n the DCT coefficients of f.

The high compression during the MPEG-related encoding process is (among other procedures) based on the quantization of the DCT coefficients, which in turn results in loss of high frequency coefficients. Within a MPEG block/macroblock, the luminance differences and discontinuities between any pair of adjacent pixels are reduced, by the encoding and compression process. On the contrary, for all the pairs of adjacent pixels, which are located across and on both edge sides of the border of adjacent DCT blocks, the luminance discontinuities are increased by the encoding process. Thus, after the quantization:

$$F_n^{'} = Q[F_n]$$

where $Q[\,]$ denotes the quantization process.

So, at the decoder side, the final reconstructed frame (after motion estimation and compensation modules) will be given by:

$$f^{'} = \sum_{n=1}^{N^2} F_n^{'} e_n$$

Thus, the perceived quality degradation per frame due to the encoding and quantization process can be expressed by an error based framework in the luminance domain Δf_Y between the original and the decoded frames.

$$\Delta f_Y \propto f_Y - f_Y^{'}$$

In this context, an average of the PQoS level for the whole encoding signal, consisting of *N* frames, can be derived by the following error-based equation:

$$< PQoS >_{video} \propto \sum_{i=1}^{N} \Delta f_{Y_i}$$

An objective perceived quality metric, which provides very reliable assessment of the video quality, based on this error-based framework, is the *SSIM* metric. The *SSIM* is a FR objective metric, which measures the structural similarity between two image/video sequences, exploiting the general principle that the main function of the human visual system is the extraction of structural information from the viewing field. Thus, considering that *f* and *f'* depicts the frames of the uncompressed and compressed signal respectively, then the *SSIM* is defined as [3, 6]:

$$SSIM(f, f') = \frac{(2\mu_f \mu_{f'} + D_1)(2\sigma_{ff'} + D_2)}{(\mu_f^2 + \mu_{f'}^2 + D_1)(\sigma_f^2 + \sigma_{f'}^2 + D_2)}$$

where μ_f, $\mu_{f'}$ are the mean of *f* and *f'*, σ_f, $\sigma_{f'}$, $\sigma_{ff'}$ are the variances of *f*, *f'* and the covariance of *f* and *f'*, respectively. The constants D_1 and D_2 are defined as:

$$D_1 = (K_1 L)^2 \quad D_2 = (K_2 L)^2$$

where *L* is the dynamic pixel range and $K_1 = 0.01$ and $K_2 = 0.03$, respectively.

Thus, *SSIM* metric can be considered as a reliable metric for quantifying PQoS for video services. Figure 2 depicts a typical example of the *SSIM* measurement per frame for the video trailer “16 Blocks”, which was encoded using the MPEG-4/H.264 standard VBR at 200 Kbps with Common Intermediate Format (CIF) resolution and 25 frames per second (fps). The instant *SSIM* vs. time curve (where time is represented by the frame sequence) varies according to the spatiotemporal activity of each frame, which causes different quality degradation for the same quantization parameters. For frames with high complexity the instant *SSIM* level drops (i.e. <0.9), while for static frames the instant PQoS is higher (i.e. >0.9 or equal to 1, which denotes no degradation at all).

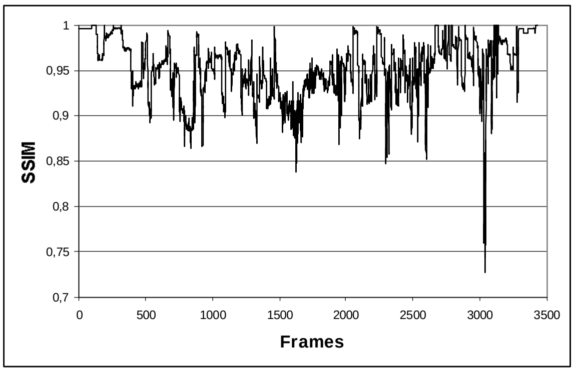

Fig. 2. The instant *SSIM* per frame of "16 Blocks" CIF 200kbps VBR

The concept of averaging the *SSIM* for the whole video duration can be exploited for deriving the mean PQoS as it was earlier defined. However, although the mean PQoS provides a single perceived quality measurement, which is more practical especially for the service providers, for long duration videos, where the spatial and temporal activity level of the content may differ significantly, the deduction of just one measurement of the perceived quality may not be accurate. In such long sequences, the proposed average metric can be combined along with a scene change detector algorithm, which will lead to calculating partial average PQoS for the various scenes. However, this case is not within the purposes of the current paper and it is not examined. The paper aims at quality issues in hand-held and small screen mobile devices, where short in duration signals are broadcasted, such as movie trailers, news or music clips with practically constant and homogeneous level of spatial and temporal activity.

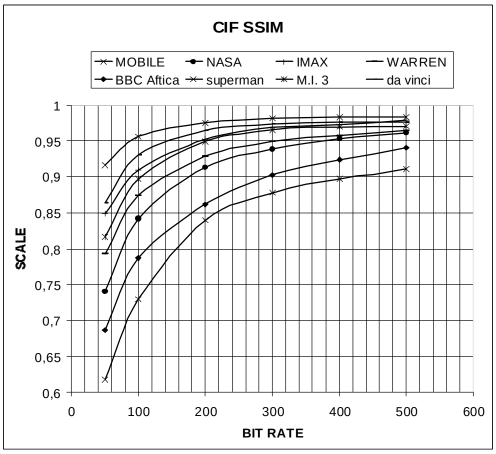

Fig. 3. The <PQoS>$_{SSIM}$ vs. bit rate curves for various test signals

In this context, eight short in duration video clips were selected and used for the needs of this paper. The experimental set consisted trailer video clips with duration up to three minutes. Each trailer clip was transcoded from its original H.264 format with Hi-Def resolution (i.e. 720p) to MPEG-4/H.264 Baseline Profile at diverse VBR bit rates. For each corresponding bit rate, a different MPEG-4/H.264 compliant file with CIF (Common Interface Format) resolution (352x288) was created. The frame rate was set at 25 frames per second (fps) for the transcoding process in all test videos.

Each encoded video clip was then used as input in the SSIM estimation algorithm. From the resulting SSIM vs. time graph (like the one in Figure 2), the <PQoS> value of each clip was calculated. This experimental procedure was repeated for each video clip in CIF resolution. The results of these experiments are depicted in Figure 3.

Referring to the curves of Figure 3, the following remarks can be made:

1. The minimum bit rate of the lowest <PQoS>$_{SSIM}$ value depends on the S-T activity level of the video clip.
2. The variation of the <PQoS>$_{SSIM}$ vs. bit rate is an increasing function, but non linear.
3. The quality improvement of an encoded video clip is not significant for bit rates higher than a specific threshold. This threshold depends on the S-T activity of the video content.

Moreover, each <PQoS>$_{SSIM}$ vs. bit rate curve can be successfully described by a logarithmic function of the general form

$$< PQoS >_{SSIM} = C_1 \ln(BitRate) + C_2$$

where C_1 and C_2 are constants strongly related to the spatial and temporal activity level of the content. Table 1 depicts the corresponding logarithmic functions for the test signals of Figure 3 along with their R^2 factor, which denotes the fitting efficiency of the theoretical curve to the experimental one.

TABLE 1. FITTING PARAMETERS AND R^2 FOR DIFFERENT VIDEO

Test Signal	Logarithmic Function	R^2 factor
Mobile	0.1295ln(x)+0.1274	0.9759
Imax	0.0563ln(x)+0.6411	0.9514
M.I. 3	0.0668ln(x)+0.5747	0.9191
Da Vinci Code	0.0474ln(x)+0.6974	0.8833
Warren	0.0738ln(x)+0.5210	0.9528
Nasa	0.0950ln(x)+0.3892	0.9595
BBC – Africa	0.1098ln(x)+0.2702	0.9875
Superman	0.0282ln(x)+0.8167	0.8859

Based on the aforementioned analysis, we can describe the derived <PQoS>$_{SSIM}$ vs. bit rate curve of each test signal with *N* total frames, which is encoded at bit rate *n* from $BitRate_{\min}$ to $BitRate_{\max}$ as a set C, where each element F_n is a triplet, consisting the <PQoS>$_{SSIM}$ of the specific bit rate, the constants C_1 and C_2, which are derived by the analytical logarithmic expression of Table 1:

$$C_{S-T} \bullet \{m : (\frac{1}{N}\sum_{i=1}^{N} SSIM(f_i), C_1, C_2)_n = F_n, n \in [BitRate_{\min}, BitRate_{\max}]\}$$

where

- *SSIM( )* is the function that calculates the perceived quality of each frame according to the *SSIM* metric

-*N* the total number of frames f_i that consists the movie *m*

Thus, deriving the sets C_n for various contents, ranging from static to very high Spatial and Temporal (S-T) ones, a reference hyper set RS, containing a range of C_{S-T} sets for specific spatiotemporal levels can be deduced:

$$RS = \{C_{S-T_{Low}}, ..., C_{S-T_{High}}\}$$

Hence, consider an unknown video clip, which is uncompressed and the broadcaster wants to predict its corresponding C_{S-T} set that better describes its perceived quality vs. bit rate curve before the encoding process at a specific quality level. Then, it is defined for all the sets C_{S-T} the Absolute Difference Value (ADV) between the first C_{S-T} triplet element (i.e. the $<PQoS>_{SSIM}$ at a specific encoding $BitRate_i$) and the experimental measurement of the average *SSIM* for the test signal at the encoding bit rate *n*, for which all the reference sets C_{S-T} have been derived, utilizing the logarithmic equations of Table 1:

$$ADV = | F_{BitRate_i} : (\frac{1}{N}\sum_{i=1}^{N} SSIM(f_i)) - F'_{BitRate_i} : (\sum_{i=1}^{N} SSIM(f'_i)) |$$

Due to the fact that the additive property is valid, it is concluded that when the *ADV* between the average *SSIM* of the reference $F_{BitRate_i}$ and experimental $F'_{BitRate_i}$ is minimum, then the set C_{S-T}, which contains the triplet element that minimizes the *ADV*, describes better the specific video. Thus, we have successfully approximated the PQoS vs. Bit rate curve of the specific video with actual cost only one testing encoding and assessment at bit rate *n*. Then the service provider can predict analytically through the logarithmic expression all the bit rates that satisfy specific perceived quality levels, without requiring any other testing encoding processes.

Moreover, the proposed technique was also tested on a set of real captured video clips, containing content with duration spanning from 2 minutes up to 10 minutes. These video clips were captured in DV PAL format from common TV programs and then transcoded to MPEG-4/H.264. Applying the proposed model and following exactly the same procedure, the worst case mean error between the experimentally and theoretically derived MPQoS curves for the twenty real captured videos was measured to be approx. 4%. A typical result of this evaluation process is depicted on figure 4, which demonstrates the fitting match between the experimentally derived curve and the predicted one.

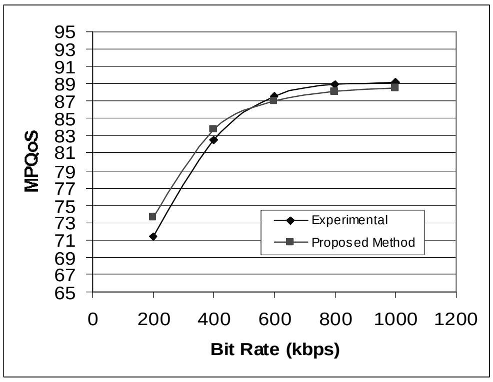

Fig. 4. Comparison of the experimentally and theoretically derived curves

Thus, one only test estimation of the average PQoS at a specific encoding bit rate is adequate for the accurate determination of the MPQoS vs. Bit Rate for a given signal.

In the next section, it is examined the case of the quality degradation during the transmission process of the broadcasting.

IV. MODELING PACKET LOSS IMPACT ON VIDEO QUALITY

In this section, we discuss the impact of the packet loss during the transmission of a video over a lossy transmission channel. Due to the fact that the frames in a MPEG video sequence are interdependent, considering a packet loss, the visual distortion caused by this packet loss will be not limited only to the frame, on which the specific lost packet belongs to. On the contrary, spatial error propagation will take place, which will infect all the frames that are dependent on the specific frame, on which the loss occurred. Thus, in order to calculate the error propagation due to a packet loss, it must be taken under consideration the interdependencies of the coded frames.

Regarding packetization schemes, all contemporary digital broadcasting systems, including the DVB and ATSC family of standards, are using the MPEG-2 Transport Stream (TS) [36] as the format of baseband data, organized in a statistically multiplexed sequence of fixed-size, 188-byte Transport Packets. Initially intended to convey MPEG-2 encoded audio and video streams, the MPEG-2 TS was eventually used also for the transport of IP traffic, with the adaptation method introduced in [37] and named as Multi Protocol Encapsulation (MPE).

Typical scenarios for fixed-size packetization schemes are a) a packet contains part of one frame, b) a packet contains the end of a frame and c) a packet contains a frame header. Independently of its content, a packet loss will create perceptual degradation and artifacts at the respective decoded frame. Therefore, the initial perceptual error will be propagated in space and time due to the interdependencies of the encoded frames and the inter-coding procedure of the motion estimation and compensation techniques.

At the user side, the PQoS degradation induced by a packet loss depends on the error concealment strategy used by the decoder. A typical concealment strategy is zero-motion concealment, in which a lost macroblock is concealed using the macroblock in the same spatial location from the closest reference frame.

Therefore, it is expected the visibility of a loss to depend on a complex interaction of its location, the video encoding parameters (i.e. GOP structure) and the underlying characteristics of the video signal itself. In this context, it is proposed a mathematical framework to model the perceptual error propagation caused by packet losses during broadcasting. More specifically, this section studies the packet loss effect on MPEG video decoding quality over error-prone broadcasting channels. We introduce an analytical model, which is used to investigate the video delivered quality and the effect of the packet loss distribution on the delivered video quality.

A. The analytical model of packet loss effect on PQoS

For evaluation purposes of the packet loss impact on the PQoS level of a broadcasting service, it is adopted an objective evaluation metric, known as Decodable Frame Rate (Q). Although the objective Q metric has been used in earlier works [38], our approach is differentiated from the existing ones because it considers the packet loss rate instead of the frame loss rate in the formula, providing a better approach for broadcasting systems. The value of Q lies between 0 and 1.0. The larger the value of Q, the better the video quality received by the end user. Where Q is defined as the fraction of decodable frame rate, which is the number of successfully decoded frames over the total number of frames sent by a video source.

$$Q = \frac{N_{dec}}{(N_{total-I} + N_{total-P} + N_{total-B})}$$

where N_{dec} is the summation of $N_{dec\text{-}I}$, $N_{dec\text{-}P}$, and $N_{dec\text{-}B}$.

Based on this modified Q metric, in the next sub-sections it is analytically calculated the expected numbers of decodable frames per type (i.e. I, B, P) based on a typical structure GOP(12,3), which is widely used in broadcasting applications for moving users due to its robustness characteristics.

In the proposed modeling, the following hypotheses are considered:

- At the decoder it is not implemented any sophisticated error concealment method.
- The decoding threshold is considered equal to one (DT=1), meaning that one packet loss causes significant quality degradation (i.e. unsuccessful decoding) of the respective frame.
- The error propagation affects all the frames that are depended on the frame, where the packet loss took place. Considering that DT=1, the dependent frames are also considered to fail during the decoding procedure.
- The packet loss rate is considered constant during the transmission of the service.

Based on these hypotheses and the modified Q metric, it is clear that the proposed approach of modelling packet loss impact on the degrading percentage of the broadcasting perceived quality of a service is an objective approach. More specifically, it is researched the degradation percentage caused by the transmission packet loss ratio in relevance to the initial quality of the video content. The relative approach provides many advances in comparison to already proposed models, namely the independence from the content dynamics, the coding standard and the structure of the packet.

Following this explanatory section, the proposed model is presented in the next sub-sections, considering constant packet loss ratio p for the whole service duration. For readability purposes, in the appendix of the paper, it can be also found the notation explanation of all the used symbols.

1) The expected number of decodable I frames ($N_{dec\text{-}I}$)

In a GOP, the I frame is successfully decodable only if all the packets that belong to the specific frame are intact received. Therefore, the probability that the I frame is successfully decodable is

$$S(I) = (1-p)^{C_I}$$

Consequently, the expected number of correctly decodable I frames for the whole video is

$$N_{dec\text{-}I} = (1-p)^{C_I} * N_{GOP}$$

2) The expected number of decodable P frames ($N_{dec\text{-}P}$)

In a GOP, P frames are successfully decodable only if the preceding I or P frames are also decodable and all the packets that belong to the P frame under examination have been successfully received. In a GOP, there are N_p P frames, and depending on their position, the probability of a P frame to be decodable is

$$S(P_1) = (1-p)^{C_I} * (1-p)^{C_P} = (1-p)^{C_I + C_P}$$

$$S(P_2) = (1-p)^{C_I} * (1-p)^{C_P} * (1-p)^{C_P} = (1-p)^{C_I + 2C_P}$$

$$\ldots\ldots$$

$$S(P_{N_p}) = (1-p)^{C_I} * (1-p)^{N_P * C_P} = (1-p)^{C_I + N_P * C_P}$$

Thus, the expected number of successfully decodable P frames for the whole video is

$$N_{dec\text{-}P} = (1-p)^{C_I} * \sum_{j=1}^{N_P} (1-p)^{jC_P} * N_{GOP}$$

3) The expected number of decodable B frames ($N_{dec\text{-}B}$)

In a GOP, B frames are decodable only if the preceding and succeeding I or P frames are both decodable and all the respective packets that consist the specific B frame have been successfully received. Considering that B frames throughout the GOP structure have the same dependencies, we examine the consecutive B frames as composing a B group, except for the last B frame in a GOP, which is dependent from the preceding P frame and succeeding I frame (making it straight forward dependent on two successive I frames). In a GOP, the probability of the B frame that is decodable is

$$S(B_1) = (1-p)^{C_I} * (1-p)^{C_P} * (1-p)^{C_B}$$

$$S(B_2) = (1-p)^{C_I} * (1-p)^{2C_P} * (1-p)^{C_B}$$

$$\ldots\ldots .$$

$$S\left(B_{\frac{N}{M}-1}\right) = (1-p)^{C_I} * (1-p)^{\left(\frac{N}{M}-1\right)*C_P} * (1-p)^{C_B}$$

$$S\left(B_{\frac{N}{M}}\right) = (1-p)^{2C_I} * (1-p)^{\left(\frac{N}{M}-1\right)*C_P} * (1-p)^{C_B}$$

Thus, the expected number of correctly decodable B frames for the whole video is

$$N_{dec\text{-}B} = (M-1) * \sum_{j=1}^{\frac{N}{M}} S(B_j) * N_{GOP}$$

$$= \left[(M-1) * (1\text{-}p)^{C_I} * \sum_{j=1}^{N_P} (1\text{-}p)^{jC_P} * (1\text{-}p)^{C_B} + (M-1) * (1\text{-}p)^{2C_I} * (1\text{-}p)^{N_P C_P} * (1\text{-}p)^{C_B}\right] * N_{GOP}$$

$$= \left[(1\text{-}p)^{C_I + N_P C_P} + \sum_{j=1}^{N_P} (1\text{-}p)^{jC_P}\right] * (M-1) * (1\text{-}p)^{C_I + C_B} * N_{GOP}$$

Based on the aforementioned proposed estimations of successfully decodable frames for each frame type, the modified Q metric becomes:

$$Q = \frac{N_{dec}}{(N_{total\text{-}I} + N_{total\text{-}P} + N_{total\text{-}B})} = \frac{N_{dec\text{-}I} + N_{dec\text{-}p} + N_{dec\text{-}B}}{(N_{total\text{-}I} + N_{total\text{-}P} + N_{total\text{-}B})} \Rightarrow$$

$$Q = \frac{(1\text{-}p)^{C_I} * N_{GOP} + (1\text{-}p)^{C_I} * \sum_{j=1}^{N_P} (1\text{-}p)^{jC_P} * N_{GOP} + \left[(1\text{-}p)^{C_I + N_P C_P} + \sum_{j=1}^{N_P} (1\text{-}p)^{jC_P}\right] * (M-1) * (1\text{-}p)^{C_I + C_B} * N_{GOP}}{(N_{total\text{-}I} + N_{total\text{-}P} + N_{total\text{-}B})}$$

Therefore, considering a transmission channel with constant packet loss ratio p, the respective Q rate of successfully decoded frames (i.e. frames without containing any perceptual degradation) can be analytically estimated. In other words, the proposed model provides a degradation parameter, which acts in a relative way to the initial quality level of the broadcasting service.

B. *Experimental Evaluation of the Proposed Model*

The proposed model of packet loss impact on the PQoS degradation of the transmitted video is experimentally evaluated considering two discrete packet loss schemes: The random uniform model, which provides the distributed losses with the mean loss rate (p) and the Gilbert-Elliot (GE) model [39], which provides for the same percentage rate, the packet losses grouped in bursts, approximating by this way the behavior of real wireless error-prone transmission channels.

For clarity purposes, Figure 5 provides a graphical representation of the used packet loss schemes.

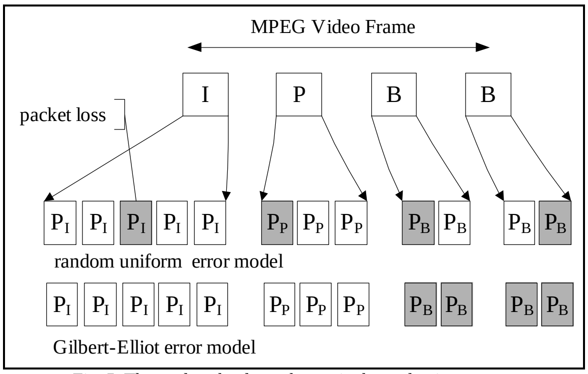

Fig. 5. The used packet loss schemes in the evaluation process

The experiments were performed on NS-2. For the evaluation purposes, the video trace “Aladdin” was selected, which is composed of 89998 video frames, including 7500 I frames, 22500 P frames, and 59998 B frames at QCIF MPEG-4/H.264 format and GOP(12,3).

TABLE 2
STATISTICS OF TEST SIGNAL ‘ALLADIN’

	Total	I frame	P frame	B frame
Number of frames	89998	7500	22500	59998
Number of packets	1086789	195010	321444	570335
C_I, C_P, C_B	N/A	26.001	14.286	9.506

Table 2 contains the statistics for the test signal, considering 188 bytes transmission, which is consistent with the MPEG-2 TS and the DVB-H standard.

For both loss schemes under test, the packet loss rate ranges from 0.02 to 0.2, considering intervals of 0.02 and transmitting packet size equal to 188 bytes.

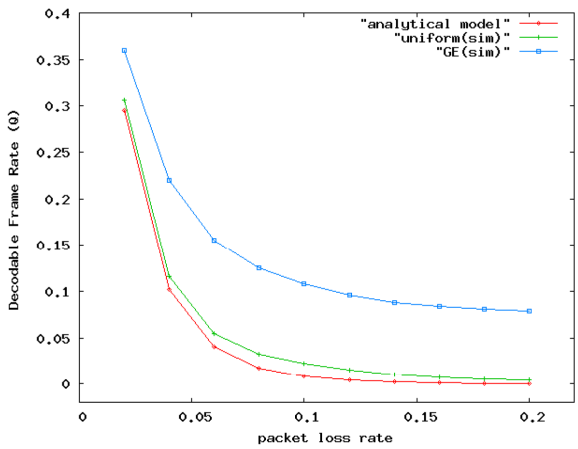

Fig. 6. Comparison of the results derived from the simulation and the analytical model.

Figure 6 shows the successfully decodable frame ratio for varying packet loss rates under the uniform and G-E packet loss schemes. Considering that the uniform distribution corresponds to the theoretically worst case scenario for decoders with DT=1 and that appears to have a significant good match between the theoretically expected video quality degradation curve and the corresponding experimentally derived one, the validity of the proposed model has been proved. For the case of G-E distribution, it is shown that the effect of burst packet losses during the transmission on the delivered video quality causes less severe degradation than the equivalent uniformly distributed case.

Moreover, in both of these models the video quality of simulation is better than analytical model. Hence, the analytical model provides the predicted bounds of the quality of the MPEG video transmission over a lossy transmission channel.

V. The Proposed End-To-End Framework

Based on the aforementioned proposed theoretical models of video quality prediction at a pre-encoding state and packet loss modeling, this section proposes an end-to-end video quality assessment framework of MPEG-based audiovisual broadcasting services for hand-held and mobile wireless broadcast systems, which is based on the combination and exploitation of the two proposed models.

For demonstrating purposes of the proposed end-to-end framework, we consider that a hypothetical Content Provider wants to broadcast a music video clip at various quality levels and possesses the reference hyper set RS, containing the C_{S-T} sets derived from the test signals of Table 1. Initially, the music clip under examination is encoded at MPEG-4/H.264 CIF 100 kbps. Then, the resulted encoded clip is used as input to the *SSIM* algorithm and the resulted instant *SSIM* curve is used for the estimation of the <*SSIM*> value, which is estimated equal to 0.8. Afterwards, using this value as input in the ADV equation, it is defined the C_{S-T} that minimizes the ADV and therefore contains the optimal triplet element for the analytical description of the signal under test. More specifically, the derived <*SSIM*> value, the optimal C_{S-T} set belongs to *BBC Africa* reference clip. Thus, the equation that describes better the variation of the $<\text{PQoS}>_{SSIM}$ vs. the bit rate is

$$<PQoS>_{SSIM} = 0.1098\ln(Bit\ Rate)+0.2702$$

Consequently, if the content provider wishes to offer this video clip at the perceptual qualities 0.70, 0.80 and 0.90, then by using the above equation is able to estimate the corresponding bit rates in a pre-encoding process. Table 3 shows the corresponding encoding bit rate values for the specific video clip.

TABLE 3
PREDICTED BIT RATE VALUES FOR SPECIFIC QUALITY LEVELS

$<\text{PQoS}>_{SSIM}$	BR (Kbps)
0.7	50.12
0.8	124.60
0.9	309.79

Afterwards, considering that a monitoring system provides the average packet loss rate at the transmission channel and it is for example 0.02, then it can be predicted from the packet loss model (see Figure 6) that the worst case degradation percentage is that the end-user will experience video quality degradation for the 70% of the total duration of the sequence. For the rest 30%, the user will watch normal playback without any perceived artifacts. Thus, if the Content Provider would like to calculate a representative value of the Expected Delivered Video Quality (EDVQ) level at the content consumer, the following equation is proposed:

$$EDVQ = (Initial\_Video\_Quality) * (Percentage\_of\_Succesfully\_Decoded\_Frames)$$

where the objective metric Q of the proposed mapping model is used as degradation multiplier to the initial perceived quality level, which has been specified pre-encodingly by the proposed prediction model. Therefore, the combination of the discrete two models provides a prediction for the worst case degradation scenario, if error concealments methods are not taken under consideration and the D.T. is considered equal to 1.0.

VI. Conclusion

This paper presents a theoretical framework for end-to-end video quality prediction for MPEG-based broadcasting services.

The proposed framework encloses two discrete models: i) a model for predicting the video quality of an encoded signal at a pre-encoding stage and ii) a model for mapping packet loss ratio of the transmitting channel to video quality degradation. The efficiency of both discrete models has been experimentally validated, proving by this way the accuracy of the proposed framework, which combines the discrete models into a common end-to-end video quality assessment framework.

The advances of the proposed framework are its generic nature, since it can be applied on MPEG-based encoded sequences, independently of the selected encoding parameters, subject to specific GOP structure. Moreover, it is also introduced the novel issue of predicting the video quality of an encoded service at a pre-encoding state, which provides new facilities at the broadcaster side. Also, by applying the randomly uniform packet loss model, the proposed framework overpasses any stochastic predicaments in mapping the packet loss ratio to video quality degradation, since it calculates and demonstrates the worst case scenario.

ACKONWLEDGEMENT

This paper is an invited extended version of the conference paper H. Koumaras, A. Kourtis, C-H Lin, C-K Shieh, "Theoretical Framework for End-to-End Video Quality Prediction of MPEG-based Sequences" published in ICNS 2007.

Part of the work in this paper has been performed within the research framework of FP7 ICT-214751 ADAMANTIUM Project.

APPENDIX

NOTATIONS USED IN THE PAPER

$N_{total\text{-}I}$ $N_{total\text{-}P}$ $N_{total\text{-}B}$	The total number of each type of frames.
$N_{dec\text{-}I}$ $N_{dec\text{-}P}$ $N_{dec\text{-}B}$	The number of decodable frames in each type.
N_{dec}	The total number of decodable frames in the video flow.
N_{GOP}	The total number of GOPs in the video flow.
C_I C_P C_B	The mean number of packets that transport the data of each frame type
p	Packet loss rate

REFERENCES

[1] Wang, Z., H.R. Sheikh, and A.C. Bovik, Objective video quality assessment, in The Handbook of Video Databases: Design and Applications, B. Furht and O. Marqure, Editors. 2003, CRC Press. p. 1041-1078.

[2] VQEG. Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment. 2000. Available : http://www.vqeg.org.

[3] Wang, Z., A.C. Bovik, and L. Lu. Why is image quality assessment so difficult? in IEEE International Conference on Acoustics, Speech, and Signal Processing. 2002.

[4] Ulrich Engelke and Hans-JÄurgen Zepernick, “Perceptual-based Quality Metrics for Image and Video Services: A Survey”, 3rd EuroNGI Conference on Next Generation Internet Networks, Trondheim, Norway, 21-23 May 2007

[5] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[6] Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," Signal Processing: Image Communication, special issue on “Objective video quality metrics”, vol. 19, no. 2, pp. 121-132, Feb. 2004.

[7] M. Ries, C. Crespi, O. Nemethova, M. Rupp, “Content Based Video Quality Estimation for H.264/AVC Video Streaming”, in Proc. Proceedings of IEEE Wireless and Communications & Networking Conference, Hong Kong, March, 2007

[8] Eric A. Silva, Karen Panetta, Sos S Agaian, “Quantifying image similarity using measure of enhancement by entropy”, Mobile Multimedia/Image Processing for Military and Security Applications 2007, Sos S. Agaian, Sabah A. Jassim, Editors, 65790U, Proceedings of SPIE -- Volume 6579, May. 2, 2007

[9] Gunawan, I.P. and M. Ghanbari. Reduced-Reference Picture Quality Estimation by Using Local Harmonic Amplitude Information. in London Communications Symposium 2003. 2003.

[10] M. Montenovo, A. Perot, M. Carli, P. Cicchetti, A. Neri, Objective evaluation of video services. Proc. of 2nd Int. Workshop on Video Processing and Quality Metrics for Consumer Electronics, 2006.

[11] S. S. Hemami, M. A. Masry, A scalable video quality metric and applications. Proc. of 1st Int. Workshop on Video Processing and Quality Metrics for Consumer Electronics, 2005.

[12] O. A. Lotfallah, M. Reisslein, S. Panchanathan, A framework for advanced video traces: Evaluating visual quality for video transmission over lossy networks. (Article ID 42083) EURASIP Journal on Applied Signal Processing, 2006. 2006.

[13] Zhou Wang, Guixing Wu, Hamid R. Sheikh, Eero P. Simoncelli, En-Hui Yang and Alan C. Bovik, Quality-Aware Images. IEEE Transactions on Image Processing.

[14] H. R. Wu, M. Yuen, A generalized block-edge impairment metric for video coding. IEEE Signal Processing Letters, 1997. 4(11): p. 317-320.

[15] P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahim, A no-reference perceptual blur metric. in Proc. of IEEE Int. Conf. on Image Processing, 2002. 3: p. 57-60.

[16] J. Caviedes, S. Gurbuz, No-reference sharpness metric based on local edge kurtosis. in Proc. of IEEE Int. Conf. on Image Processing, 2002. 3: p. 53-56.

[17] A. Cavallaro, S. Winkler, Segmentation-driven perceptual quality metrics. in Proc. of IEEE Int. Conf. on Image Processing, 2004. 5: p. 3543-3546.

[18] R. R. Pastrana-Vidal, J. C. Gicquel, Automatic quality assessment of video fluidity impairments using a no-reference metric. in Proc. of 2nd Int. Workshop on Video Processing and Quality Metrics for Consumer Electronics, 2006.

[19] M. C. Q. Farias, S. K. Mitra, No-reference video quality metric based on artifact measurements. in Proc. of IEEE Int. Conf. on Image Processing, 2002. 3: p. 141-144.

[20] X. Marichal, W. Y. Ma, H. J. Zhang, Blur determination in the compressed domain using DCT information. in Proc. of IEEE Int. Conf. on Image Processing, 2002. 2: p. 386-390.

[21] R. Ferzli, L. J. Karam, J. Caviedes, A robust image sharpness metric based on kurtosis measurement of wavelet coefficients,. Proc. of 1st Int. Workshop on Video Processing and Quality Metrics for Consumer Electronics, 2005.

[22] X. Marichal, W. Y. Ma, H. J. Zhang, Blur determination in the compressed domain using DCT information. in Proc. of IEEE Int. Conf. on Image Processing, 2002. 2: p. 386-390.

[23] S. Liu, A. C. Bovik, Effcient dct-domain blind measurement and reduction of blocking artifacts. IEEE Transactions on Circuits and Systems for Video Technology, 2002. 12(12): p. 1139-1149.

[24] M. Ries, O. Nemethova, M. Rupp, Reference-free video quality metric for mobile streaming applications. in Proc. of 8th Int. Symp. on DSP and Communication Systems & 4th Workshop on the Internet, Telecommunications and Signal Processing, 2005: p. 98-103.

[25] Lu, L., et al. Full-reference video quality assessment considering structural distortion and no-reference quality evaluation of MPEG video. in IEEE International Conference on Multimedia. 2002.

[26] H. Koumaras, A. Kourtis, D. Martakos, “Evaluation of Video Quality Based on Objectively Estimated Metric”, Journal of Communications and Networking, Korean Institute of Communications Sciences (KICS), Vol. 7, No.3, pp.235-242, Sep 2005.

[27] H. Koumaras, A. Kourtis, D. Martakos, J. Lauterjung, “Quantified PQoS Assessment Based on Fast Estimation of the Spatial and Temporal Activity Level”, Multimedia Tools and Applications, Springer Editions Vol. 34(3), pp. 355-374, September 2007.

[28] H. Koumaras, E. Pallis, G. Xilouris, A. Kourtis, D. Martakos, J. Lauterjung, “Pre-Encoding PQoS Assessment Method for Optimized Resource Utilization”, 2nd Inter. Conference on Performance Modelling and Evaluation of Heterogeneous Networks, Het-NeTs04, Ilkley, United Kingdom, 2004.

[29] S. Kanumuri, P. C. Cosman, A.R. Reibman, V.A. Vaishampayan, “Modeling Packet-Loss Visibility in MPEG-2 Video”, IEEE transactions on Multimedia, Vol.8, No.2, pp.341-355, April 2006.

[30] Z. He, H. Xong, “Transmission Distortion Analysis for Real-Time Video Encoding and Streaming over Wireless Networks”, IEEE Transactions on Circuits and Systems for Video Technology, Vol.16, No.9, pp.1051-1062, September 2006

[31] J. Mitchell and W. Pennebaker. MPEG Video: Compression Standard. Chapman and Hall, 1996. ISBN 0412087715

[32] Cheng-Han Lin, Chih-Heng Ke, Ce-Kuen Shieh, Naveen Chilamkurti, “The Packet Loss Effect on MPEG Video Transmission in Wireless Networks”, The IEEE 20th International Conference on Advanced Information Networking and Applications (AINA'06), April 18-20, 2006, Vienna, Austria

[33] A. Ziviani, B. E. Wolfinger, J. F. Rezende, O. C. M. B. Duarte, and S. Fdida, "*Joint Adoption of QoS Schemes for MPEG Streams,*" Multimedia Tools and Applications Journal, to appear.
[34] J. P.Ebert, A.Willig, *A Gilbert-Elliot Bit Error Model and the Efficient Use in Packet Level Simulation*, Technical Repoert, TKN-99-002, Technical University of Berlin, March 1999.
[35] C.-H.-Ke, C.-H.-Lin, C.-K. Shieh, and W.-S. Hwang, "A Novel Realistic Simulation Tool for Video Transmission over Wireless Network," presented at The IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC2006), Taiwan, 2006.
[36] ISO/IEC 13818-1, Generic Coding of Moving Pictures and Associated Audio Information (MPEG-2) Part 1: Systems, 1996
[37] ETSI EN 301 192, Digital Video Broadcasting (DVB): DVB Specification for data broadcasting, European Standard, v.1.4.1, Nov.2004
[38] A. Ziviani, B. E. Wolfinger, J. F. Rezende, O. C. M. B. Duarte, and S. Fdida, "Joint Adoption of QoS Schemes for MPEG Streams," Multimedia Tools and Applications Journal, vol. 26, no. 1, pp. 59-80, May 2005.
[39] J. P.Ebert, A.Willig, A Gilbert-Elliot Bit Error Model and the Efficient Use in Packet Level Simulation, Technical Repoert, TKN-99-002, Technical University of Berlin, March 1999.
[40] H. Koumaras, A. Kourtis, C-H Lin, C-K Shieh, A Theoretical Framework for End-to-End Video Quality Prediction of MPEG-based Sequences, Third International Conference on Networking and Services ICNS07, 19-25 June 2007 Page(s):62 – 62, Athens, Greece 2007.

Harilaos Koumaras was born in Athens, Greece in 1980. He received his BSc degree in Physics in 2002 from the University of Athens, Physics Department, his MSc in Electronic Automation and Information Systems in 2004, being scholar of the non-profit organization Alexander S Onassis, from the University of Athens, Physics and Informatics Department and his PhD in 2007 on digital video quality prediction from the University of Athens, Informatics Department, having granted the four-year scholarship of NCSR "Demokritos". He has received twice the Greek State Foundations (IKY) scholarship during the academic years 2000-01 and 2003-04. He has also granted with honors the classical piano and harmony degrees from the classical music department of Attiko Conservatory. He joined the Digital Telecommunications Lab at the National Centre of Scientific Research "Demokritos" in 2003 and since then he has participated in EU-funded and national funded projects with presentations and publications at international conferences, scientific journals and book chapters. At the same time, he is an associate lecturer at the Business College of Athens (BCA) and City University of Seattle, teaching modules related to Information Technology, Data Networks and Mathematics. His research interests include objective/subjective evaluation of the perceived quality of multimedia services, video quality and picture quality evaluation, video traffic modeling, digital terrestrial television and video compression techniques. Currently, he is the author or co-author of more than 30 scientific papers in international journals, technical books and book chapters, numbering 41 non-self citations. He is an editorial board member of Telecommunications Systems Journal and a reviewer of EURASIP Journal of Applied Signal Processing and IEEE Transactions on Broadcasting. Dr. Koumaras is a member of IEEE, SPIE and National Geographic Society.

Cheng-Han Lin is currently a Ph.D. candidate studying in the Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan. Lin received his MS and BS degree from the Electrical Engineering Department of National Chung Cheng University in 2002 and 2004. His current research interests include wireless MAC protocols, multimedia communications, and QoS network.

Ce-Kuen Shieh is currently a professor teaching in the Department of Electrical Engineering, National Cheng Kung University. He received his PhD, MS, and BS degrees from the Electrical Engineering Department of National Cheng Kung University, Tainan, Taiwan. His current research areas include distributed and parallel processing systems, computer networking, and operating systems.

Anastasios Kourtis received his B.S. degree in Physics in 1978 and his Ph.D. in Telecommunications in 1984, both from the University of Athens. Since 1986, he has been a researcher in the Institute of Informatics and Telecommunications of the National Centre for Scientific Research "Demokritos", currently ranking as Senior Researcher. His current research activities include, digital terrestrial interactive television, broadband wireless networks, Perceived Quality of video services, end to end QoS and real time bandwidth management in satellite communications. He is author or co-author of more than 80 scientific publications in international scientific journals, edited books and conference proceedings. Dr. Kourtis has a leading participation in many European Union funded research projects in the frame of IST/FP5/FP6 (MAMBO, SOQUET, CREDO, WIN, LIAISON, ENTHRONE). He has also coordinated three European funded Specific Targeted Research Projects (REPOSIT, ATHENA, IMOSAN).