Objective and Subjective Assessment Methods of Video Quality in Multimedia Broadcasting

Harilaos G. Koumaras

koumaras@iit.demokritos.gr

NCSR Demokritos, Institute of Informatics and Telecommunications, Athens, Greece

1. Introduction

Current digital broadcasting networks throughput rates are insufficient to handle raw video data in real time, even if low spatial and temporal resolution (i.e. frame size and frame rate) has been selected. Towards alleviating the network bandwidth requirements for efficient transmission of audiovisual content, coding/compression techniques have been applied on raw video data, performing compression on both temporal and spatial redundancy of the content.

More specifically, coding applications that are specialized and adapted in broadcasting digitally encoded audiovisual content have known an explosive growth in terms of development, deployment and provision. Video coding is defined as the process of compressing and decompressing a digital video sequence, which results in lower data volumes, besides enabling the transmission of video signals over bandwidth-limited means, where uncompressed video signals would not be possible to be transmitted.

In this multi-evolutionary environment, the new era of digital video broadcasting has arrived and the beyond typical analog-based transmission for broadcasting services is a fact, setting new research challenges for the

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

assessment of Perceived Quality of Service (PQoS) under the latest video encoding and broadcasting standards.

The majority of the compression standards have been proposed by the International Telecommunication Union (ITU) and the International Organization for Standardization (ISO) bodies, by introducing the following standards H.261, H.263, H.263+, H.263++, H.264, MPEG-1, MPEG-2, MPEG-4 and MPEG-4 Advanced Video Coding (AVC). Some of the aforementioned standards were developed in partnership of ITU with Moving Pictures Expert Group (MPEG), under the venture name Joint Video Team (JVT), exploiting similar coding techniques developed by each body separately.

Each standard was designed and targeted a specific service and application, featuring therefore specific parameters and characteristics. For example MPEG-1 was proposed by MPEG in order to be used by the Video Compact Disc (VCD) medium, which stores digital video on a Compact Disc (CD) with a quality almost similar to that of an analog VHS video. In 1994 MPEG-2 was proposed for encoding audiovisual content for broadcast signals, exploiting interlace format. MPEG-2 is also the coding format used by the widely successful commercial Digital Versatile Disc (DVD) medium, while the latest H.264, or MPEG-4 Part 10 AVC aims at providing high broadcasting video quality at very low bit rates on a wide variety of applications, networks and systems or High Definition resolution through Blu-Ray Discs and Multimedia content.

Video compression standards exploit in their algorithms the high similarity of the depicted data in the spatial, temporal and frequency domain within and between subsequent frames of a video sequence [1]. Correlating this redundancy in these three domains, it is achieved data compression against a certain amount of visual data loss, which from the one hand cannot be retrieved but on the other hand it is not perceived by the viewers, since it is not 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 at the cost of a partial loss of the data and subsequently video quality degradation of the initial signal. Due to the fact that the parameters with strong influence on the video quality are normally those which play the most important role in the bit rate during the encoding/compression process, the issues of the user satisfaction and video quality assessment in correlation with the encoding parameters have been raised.

One of the future visions is the provision of audiovisual content at various quality and price levels. There are many approaches to this issue, one being the Perceived Quality of Service (PQoS) concept. The evaluation of the PQoS for audiovisual content will provide a user with a range of poten-

Author’s Accepted Manuscript (AAM) Version: Chapter contribution in “*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*“, Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

tial choices, covering the possibilities of low, medium or high quality levels. Moreover the PQoS evaluation gives the service provider and network operator the capability to minimize the storage and network resources by allocating only the resources that are sufficient to maintain a specific level of user satisfaction.

The evaluation of the PQoS is a matter of procedures, each time taking place after the encoding process (post-encoding evaluation). The methods and techniques that have been proposed in the bibliography mainly aim at:

- Determining the encoding settings (i.e. resolution, frame rate, bit rate) that are required in order to carry out successfully a communication task of a multimedia application (i.e. video conference).
- Evaluating the quality level of a media clip based on the detection of artefacts on the signal caused by the encoding process.

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 performed by recurring subjective or objective video quality assessments, each time taking place after the encoding process (repetitive post-encoding evaluations). However, 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 artefact on the decoded frame and therefore quality degradation.

In this context, this chapter discusses the various PQoS-related aspects that are involved in the end-to-end video quality assessment of MPEG-based broadcasting services, focusing on:

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

i. The assessment methods of the encoded broadcasting service, which aim at specifying a specific video quality level in terms of encoding bit rate and content dynamics.

ii. The impact of transmission impairments, such as the packet loss ratio during the transmission, on the respective cross QoS-related layers and delivered video quality level respectively.

The rest of the chapter is organized as follows: The next section presents and discusses both on the subjective and objective video quality assessment methods that have been proposed and are applied on digitally encoded and compressed video signals. Afterwards, it is discussed the impact of various transmission conditions and impairments, such as packet loss occurred by bad transmission conditions, on the deduced perceptual quality. In this context, it is described how this impact can be modelled in a deterministic way across the various QoS layers of the broadcasting service in terms of video quality degradation over error-prone transmission channel. Finally, the last section concludes the chapter.

2. Video Quality Assessment Methods at the Encoding/Generation Phase

The advent of quality evaluation was the application of pure error-based sensitive framework between the encoding and the original/uncompressed video sequence. These primitive methods, although they provided a quantitative approach of the quality degradation between the encoded and original signals, they did not provide reliable measurements of the perceived quality, because they miss out the characteristics and sensitivities of the Human Visual System. In this context, the most widely used primitive methods and quality metrics that used an Error Sensitivity framework are the Peak Signal to Noise Ratio (PSNR) and the Mean Square Error (MSE). Over the last years, emphasis has been put on developing methods and techniques for evaluating the perceived quality of digital video content mainly during the encoding process. These methods are categorized into two classes [2] : The subjective and objective ones.

- The subjective test methods involve an audience of people, who watch a video sequence and score its quality as perceived by them, under specific and controlled watching conditions.

The subjective assessment methods are further classified into classes depending on the test procedure, which may include simultaneous viewing of both degraded and original video signal (double stimulus methods) or of only one signal at a time (single stimulus methods).

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

- The objective evaluation methods, which successfully emulate the results that are derived from subjective quality assessments, based on criteria and metrics that can be measured objectively.

The objective assessment methods are further classified into classes, according to the procedure of the quality evaluation, depending on the requirement of the initial uncompressed and non-degraded content into the evaluation process. Based on this categorization, the assessment methods are named as Full Reference, Reduced Reference or No-Reference methods, representative of the requirement or not of the initial uncompressed signal. The next sections discuss and present the most popular subjective and objective assessment method categories.

2.1 Subjective Assessment Methods

The subjective test methods have been mainly proposed by International Telecommunications Union (ITU) and Video Quality Experts Group (VQEG), involving an audience of people, who watch and score the quality of a video sequence as perceived by them, under specific and controlled watching conditions. Afterwards, the statistical analysis of the collected data is used for the evaluation of the perceived quality, usually exploiting the Mean Opinion Score (MOS) as the most reliable and typical metric of quality measurement.

Subjective test methods are described in ITU-R Rec. BT.500-11 [3] and ITU-T Rec. P.910 [4], suggesting specific viewing conditions, criteria for observers and test material selection, assessment procedure description and statistical analysis methods. The ITU-R Rec. BT.500-11 described subjective methods that are specialized for television applications, whereas ITU-T Rec. P.910 is intended for multimedia applications.

The subjective methods depending on the number of the simultaneous sequences under test are mainly classified as single or double stimulus when one or two signals are used respectively [5].

In the Single Stimulus Methods multiple separate scenes are shown simultaneously and the viewers are asked to evaluate each one separately. Depending on the playback order of the test signals, there SS methods are classified to two approaches: SS with no repetition of test scenes and SS where the test scenes are repeated multiple times. Three different scoring methods are used: Adjectival, Numerical and Non-categorical (i.e. a continuous scale with no numbers). Representative single stimulus methods are the following:

- Single Stimulus Method (SSM)

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

- Absolute Category Rating (ACR)
- Single Stimulus Continuous Quality Evaluation (SSCQE)

In the Double Stimulus Methods, the observers watch multiple references and degraded scene pairs. The order of the reference scene relative to the degraded one may differ depending on the implemented method. Also the viewers may not be aware of which signal is the reference and/or the degraded one. Scoring is usually on an overall impression scale of impairment either using adjectival or non-categorical scale. Representative double stimulus methods are the following:

- Double Stimulus Continuous Quality Evaluation (DSCQE)
- Double Stimulus Impairment Scale (DSIS)
- Degradation Category Rating (DCR)
- Pair Comparison method (PC)

2.2 Objective Assessment Methods

The preparation and execution of subjective tests is costly and time consuming and its implementation today is mainly limited to scientific purposes, especially at Video Quality Experts Group (VQEG) experiments.

For this reason, a lot of effort has recently been focused on developing cheaper, faster and easier applicable objective evaluation methods. These techniques successfully emulate the subjective quality assessment results, based on criteria and metrics that can be measured objectively. The objective methods are classified, according to the availability of the original video signal, which is considered to be in high quality.

The majority of the proposed objective methods in the literature require the undistorted source video sequence as a reference entity in the quality evaluation process, and due to this are characterized as Full Reference Methods [6]. The methods perform multiple channel decomposition of the video signal, where the proposed objective method is applied on each channel, which features a different weigh factor according to the characteristics of the Human Visual System. The basic block diagram of the full reference methods with multiple channels is depicted on Figure 1. These methods emulate characteristics of the Human Visual System (HVS) using Contrast Sensitivity Functions (CSF), Channel Decomposition, Error

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

Normalization, Weighting and finally Minkowski error pooling for combining the error measurements into single perceived quality estimation [7].

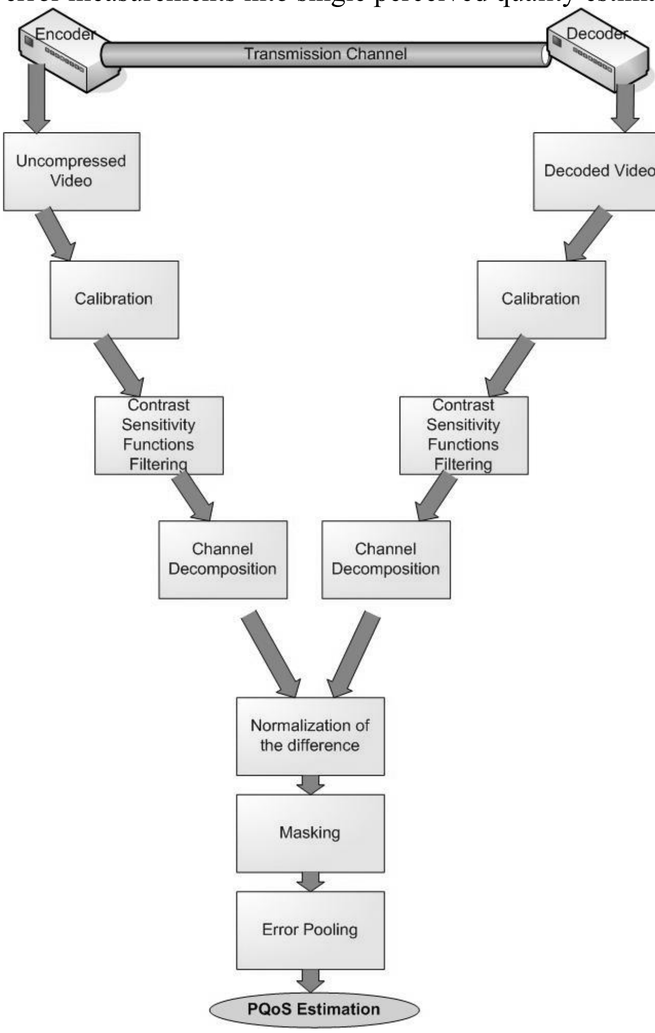

Figure 1. Full Reference Methods with multiple channels

Similarly, in the bibliography it has been proposed some full reference methods of single channel, where the proposed objective metric is applied homogeneously on the video signal, without considering varying weight functions. The block diagram of these methods is depicted on Figure 2. However it has been reported [8, 9] that these complicated methods do not provide more accurate results than the simple mathematical measures (such as PSNR). Due to this some new full reference metrics that are based

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

on the video structural distortion, and not on error measurement, have been proposed [7, 10-13].

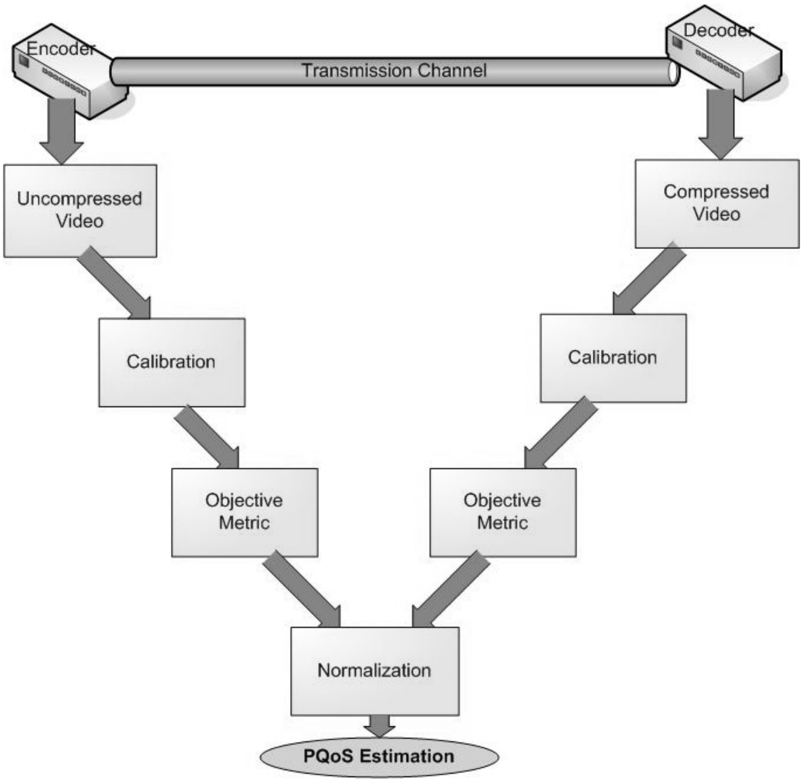

Figure 2. Full Reference Methods with single channel

On the other hand, the fact that these methods require the original video signal as reference deprives their use in commercial video service applications, where the initial undistorted clips are not always accessible. Moreover, even if the reference clip is available, then synchronization predicaments between the undistorted and the distorted signal (which may have experienced frame loss) make the implementation of the Full Reference Methods difficult and impractical.

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) [14, 15]. The block diagram of the reduced reference methods is depicted on Figure 3.

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

Finally, some methods and techniques have been proposed in the bibliography that do not require any reference video signal (No Reference Methods) [16, 17].

Nevertheless, due to the fact that the future vision is the provision of audiovisual content at various quality and price levels [18], there is great need for developing methods and tools that will help service providers to predict quickly and easily the PQoS level of a media clip. These methods will enable the determination of the specific encoding parameters that will satisfy a certain quality level. All the previously mentioned post-encoding methods may require repeating tests in order to determine the encoding parameters that satisfy a specific level of user satisfaction. This procedure is time consuming, complex and impractical for implementation on the broadcasting services.

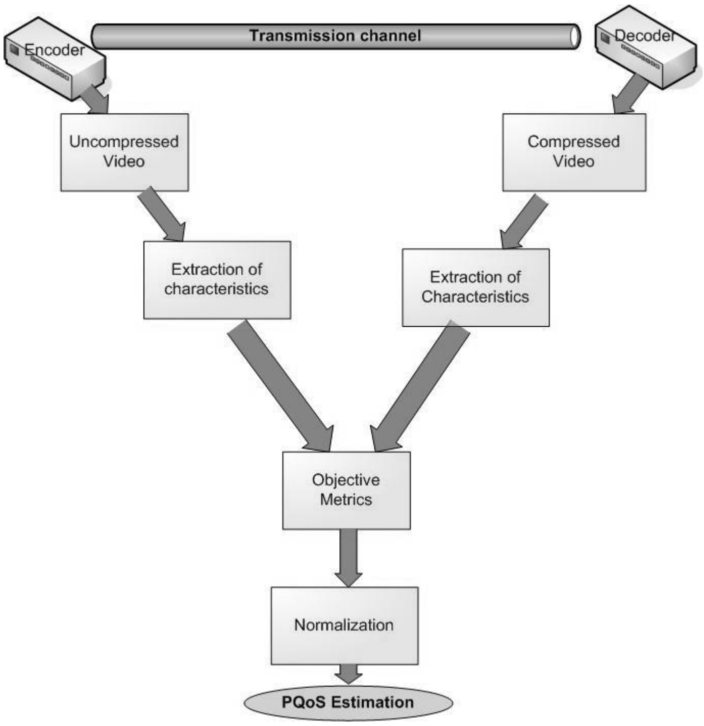

Figure 3. Reduced Reference Methods

Towards this, recently it has been performed research in the field of pre-encoding estimation and prediction of the PQoS level of a multimedia service as a function of the selected resolution and the encoding bit rate [19-24]. These methods provide fast and quantified estimation of the PQoS,

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

taking into account the instant PQoS variation due to the Spatial and Temporal activity within a given encoded sequence. Quantifying this variation by the Mean PQoS (MPQoS) as a function of the video encoding rate and the picture resolution, it is finally used the MPQoS as a metric for pre-encoding PQoS assessment based on the fast estimation of the spatiotemporal activity level of a video signal.

3. Video Quality Issues during Service Transmission: Translation between PQoS, AppQoS and NQoS

This section discusses how the transmission errors and impairments of the transmission channel are mapped to the various QoS-related layers of the broadcasting service. More specifically the following sub-sections refer to the Perceived QoS (PQoS), Application QoS (AppQoS) and Network QoS (NQoS) layers, discussing the various aspects of their cross mapping.

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) 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.

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, [25] 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.

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

Similarly, in [26] 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.

Regarding the mapping between the various discrete QoS layer (i.e. PQoS/AppQoS/NQoS), Table 1 defines the representative metrics of each level, which must be used and considered into any relative mapping process or model:

Table 1: Metrics of each QoS Level

Service QoS Level	Application QoS Level	Network QoS Level
User Satisfaction PQoS level Terminal Specifications	Decodable Frame Rate Decoding Threshold Encoding Parameters	Packet Loss Ratio Packet Loss Scheme Packet Size

At the service QoS level, the critical metric is the user satisfaction (i.e. PQoS level). The evaluation of the PQoS for audiovisual content will provide a user with a range of potential choices, covering the possibilities of low, medium or high quality levels (i.e. gold, silver and bronze services). Moreover the PQoS evaluation gives the service provider and network operator the capability to minimize the storage and network resources by allocating only the resources that are sufficient to maintain a specific level of user satisfaction.

As it has been already mentioned and explained in the previous sections, the evaluation of the PQoS is a matter of objective and subjective evaluation procedures, each time taking place after the encoding process (post-encoding evaluation). Subjective quality evaluation processing of video

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

streams (PQoS evaluation) requires large amount of human resources, establishing it as a time-consuming process (e.g. large audiences evaluating video/audio sequences). Objective evaluation methods, on the other hand, can provide PQoS evaluation results faster, but require large amount of machine resources and sophisticated apparatus configurations. Towards this, objective evaluation methods are based and make use of multiple metrics, which are related to the content's artifacts (i.e. tilling, blurriness, error blocks, etc.) resulting from the quality degradation due to the encoding process (see Figure).

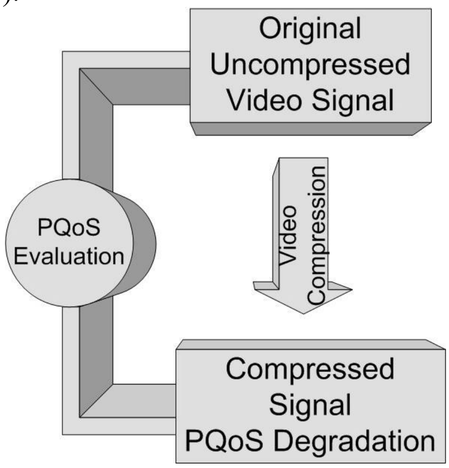

Figure 4: Concept of the PQoS evaluation

At the AppQoS level, given that during the encoding process quality degradation of the initial video content (see Figure 4) is incurred, the values of the encoding parameters (i.e. bit rate, resolution) play a major role in the resulting PQoS. Thus, the various encoding parameters must be used as metrics in quantifying the deduced PQoS level. If we also consider additional degradation due to transmission problems (i.e. limited bandwidth, network congestion), which finally result in packet loss at the video packet receiver during the service transmission, then the Decodable Frame Rate can be considered as a metric for quantifying this phenomenon. The Decodable Frame Rate Q is an application-level metric, with values ranging from 0 to 1.0. The larger the value of Q, the higher the successful decoding rate at the end-user. Q is defined as the fraction of decodable frame rate, which is the number of decodable frames (i.e. frames that are theoretically able to be decoded without considering the post-filtering or error concealment abilities of each decoder) over the total number of frames sent by a video source.

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

Since different codec and transmission techniques have different tolerance to packet loss, the theoretically expected decoding threshold will be also used as a metric in order to define the impact of the packet loss ratio on the frame loss ratio. A theoretical decoding threshold equal to 1.0 means that only one packet loss results in unsuccessful decoding of the corresponding frame, to which the missing packet is a part of.

Finally, at the NQoS level the metrics Packet Loss Ratio, Packet Loss scheme and Packet Size may be considered as key parameters. Although, it is obvious that other network statistics and phenomena may be present over a broadcasting network (e.g. jitter, delay), however all these parameters are quantified into the packet loss effect, since this is the final outcome of all these network QoS-sensitive parameters at the video packet receiver. Otherwise, if no packet loss occurs due to these phenomena, then sophisticated buffer techniques may eliminate their impact. Thus, with the appropriate approach the packet loss ratio can be considered as adequate parameter and used as a network metric to the PQoS-NQoS and NQoS-PQoS mapping. Regarding the various packet loss schemes (e.g. unified, bursty etc.), due to the stochastic nature of the PQoS degradation over an error-prone broadcasting channel, for reference purposes focus must be given on identifying the packet loss scheme, which provides the worst case scenario in terms of affecting the decodable frame rate (i.e. the delivered PQoS level) for specific packet loss ratio.

More detailed explanation of each metric and description of its scope and role is presented in the following sub-sections following two discrete directions from PQoS down to NQoS and the opposite one from the NQoS up to the PQoS.

3.1 PQoS to AppQoS and NQoS mapping

The mapping of the PQoS to the AppQoS covers the relationship between the service and the application level. Based on a predefined perceptual quality at PQoS, then the appropriate parameters at the application level (frame rate, bit rate, codec) are determined. The mapping is based on empirical data that are derived from subjective or objective quality assessments for different genres of content.

Concerning the initial preparation of the content at the requested/targeted PQoS level, a method for mapping the content dynamics/genre of the video to the encoding parameters that satisfy the requested/specific level of user satisfaction is necessary. Taking into account the instant PQoS variation due to the spatiotemporal activity within a given MPEG encoded content, the respective Mean PQoS (MPQoS) as a function of the

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

video encoding rate can be exploited as a metric for objective video quality assessment. Based on the proposed metric, this task will derive a reference table containing the encoding bit rate that satisfies specific quality (i.e. PQoS) levels depending on the spatiotemporal activity of the requested content.

Towards the specification of these reference MPQoS vs. bitrate rules, an objective quality meter tool of the PQoS level may be used, providing objective PQoS assessment for each frame within a video clip. The graphical representation of these results vs. time, demonstrates the instant PQoS of each frame within the video clip, besides indicating the Mean PQoS (MPQoS) of the entire video (for the whole clip duration). Similar experiments will be conducted for the MPQoS calculation of the same video content, each time applying different encoding parameters. The results of these experiments will be used to draw-up experimental curves of the MPQoS of the given video content, as a function of the encoding parameters. The same procedure will be repeated for a set of video sequences, each one with different spatiotemporal activity level.

More specifically, considering three discrete spatiotemporal categories (i.e. high, medium, low) and their respective MPQoS vs bitrate equations, the service provider should be able to specify the bit rate that satisfies a specific perceptual quality level. Figure 5 depicts the concept of this approach and the expected form of the PQoS vs. Bit Rate dependence as it has been reported in the relative literature [21, 27] due to the logarithmic sensitivity of the Human Visual System - HVS.

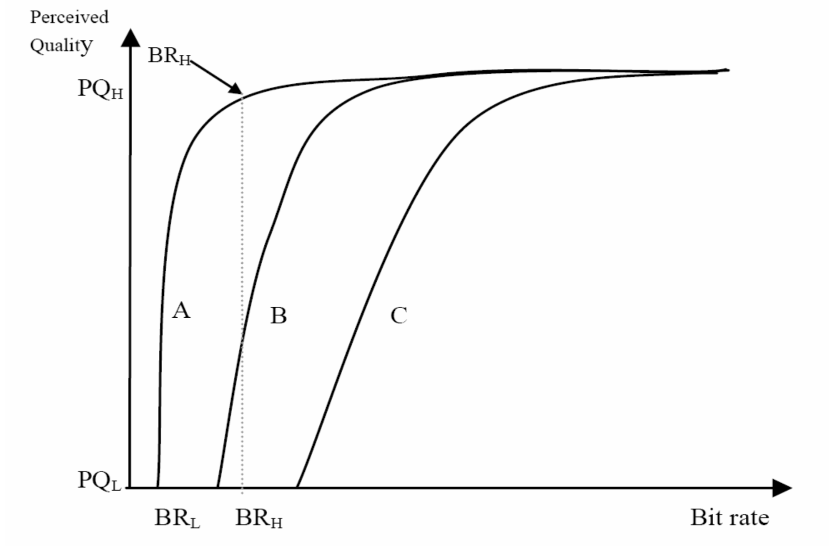

Figure 5: PQoS vs. bitrate curves for various spatiotemporal contents

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

As depicted in Figure 5, 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 (a music video clip). 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 end user (i.e. bronze service), (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) (i.e. gold service, and (c) the shape of the curve, which is defined by the content dynamics. These parameters can be experimentally derived for reference purposes and further used for defining a generic equation for describing the MPQoS vs bitrate curves. Based on relative published research [21], the respective MPQoS vs. Bit Rate curves are successfully modeled as follows:

$$MPQoS = [PQ_H - PQ_L]\,(1 - e^{-\alpha\,[BR-BR_L]}) + PQ_L,\ \alpha>0 \text{ and } BR>BR_L$$

where the parameter α is the time constant of the exponential function, which determines the shape of the curve and BR the encoding Bit Rate of the service.

The pre-encoding PQoS assessment nature of the described procedure alleviates both the machine resource requirements and the time consumption of the already existing post encoding video quality assessment methods, making PQoS evaluation quick, easy and economically affordable for commercial implementations.

The mapping of AppQoS to NQoS deals with the translation of application level parameters to parameters of the underlying network level. Multimedia services, especially broadcasting applications, tend to impose great demands on the communication networks concerning bandwidth, maximum tolerable delay, jitter, and packet loss. A pessimistic estimation of the required network resources might lead to over-provisioning of resources for a single multimedia service, resulting in bad link utilization and a waste of network resources. On the other hand, a too optimistic mapping bears the risk of congestion within the network resulting in packet loss that decreases the End-to-End QoS. Therefore, a trade-off between these extremes has to be envisaged.

In this context, the most relevant parameter at the upper application layer with direct impact to the NQoS is the video bit rate. As it has been already mentioned earlier, the selection of the appropriate video bit rate is

Author’s Accepted Manuscript (AAM) Version: Chapter contribution in “*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*“, Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

influenced by a variety of factors, such as the content dynamics, the video codec and the fidelity of the encoding.

Based on the selected video bit rate at the application level, the required network bandwidth can be derived, based on the overhead that is introduced by the protocol stack. In a typical digital broadcasting scenario, a considerable amount of headers is added to the actual media payload. This results in an overhead for each MPEG transport stream packet that is transmitted over the network.

3.2 NQoS to AppQoS and PQoS mapping

Concerning the mapping of the Network QoS sensitive parameters (like delay, packet loss etc.) to perceived video quality (i.e. PQoS) some approaches have already been proposed in the literature, which perform a very analytical statistical model of the packet-loss visual impact on the decoding video quality for MPEG-2 video sequences, specifying the various factors that affect 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). Similarly, a transmission/distortion modeling for real-time video streaming over error-prone wireless networks has also been presented, where a modeling of the impulse transmission distortion (i.e. the visual fading behavior of the transmission errors) is performed.

However, all the already proposed models are very codec and content specific, while they do not also provide any end-to-end video quality estimation, namely the degradation during the encoding process and the transmission/streaming procedure. In this framework, once the content has been prepared for delivery at the requested PQoS level, according to the reported monitored network conditions (e.g. Packet Loss rate) and the reference look up tables/rules of the NQoS to PQoS mapping, the worst case degradation to the application QoS (i.e. Undecoded or Lost Frames) and to the Service QoS (Percentage of the total duration for which the end-user will experience degraded PQoS – i.e. Delivered PQoS<Requested PQoS) could be able to predicted. According to this representation, a pre-provision assessment of the end-to-end PQoS degradation will be performed. Thus, it is necessary to develop mapping rules between the Application Level (e.g. decodable frame rate) and the Network Level (i.e.. Packet Loss, Packet Size, Packet Loss Scheme) parameters. Of course, the decodable frame rate metric of the application level may be extended as a metric to the Service Level, representing the duration percentage of the requested or no-degraded PQoS level.

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

It must be noted that during this whole mapping procedure, the sophisticated delay and delay variation phenomena may be not taken under consideration, since it can be supposed that they are successfully managed by efficient play-out buffer structures or they will eventually result in packet loss. So, the ultimate goal of the decision-taking is to find parameters for both the application and network levels, which do not violate the constraints that were imposed at the service level in terms of PQoS.

More specifically, regarding the application QoS and network QoS mapping, the translations between network packet loss ratio and Decodable Frame Rate (Q), as well as packet size and Decodable Frame Rate can be exploited. Q is an application-level metric, with values ranging from 0 to 1.0. The larger the value of Q, the higher the successful decoding rate at the end-user. Q is defined as the fraction of decodable frame rate, which is the number of decodable 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 Ndec is the sum of number of theoretically expected to be successfully decoded I, P, B frames i.e., N_{dec-I}, N_{dec-P}, and N_{dec-B}, without taking under consideration the post-filtering and error concealment techniques of the codec.

Due to the fact that the frames in a MPEG video sequence are interdependent, considering a packet loss, the visual distortion due to this packet loss will not be 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 interdependent to the specific frame, on which a packet loss occurred. Thus, in order to calculate the theoretically expected error propagation due to a packet loss, one must take under consideration the impulse transmission of the distortion.

From the hierarchical structure of MPEG encoding stream, a video frame may be considered theoretically undecodable directly or indirectly:

- Directly undecodable when the packet loss occurred in a group of packets that carry the data of the specific frame.
- Indirectly undecodable, when the packet loss occurred in a group of packets that carry the data of another frame, from which the current frame is directly depended and its successful decoding depends on the successful decoding of the corrupted frame.

For specifying the theoretically worst case scenario in order to avoid the stochastic nature of the packet loss effect on the PQoS degradation, we

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

will not consider any concealment method. So the Decodable Threshold (DT) is 1.0 (i.e. even one direct or indirect packet loss causes an undecodable frame). Therefore, our analysis will provide the worst-case of video transmission quality degradation.

Due to the very specific structure of an MPEG stream (i.e. GOP type), which is specified by successive I, P and B frames, given the deterministic fact that one packet loss results in the corresponding frame loss (i.e. DT=1.0), then the whole mapping between AppQoS and NQoS can be mathematically modeled. More specifically, given a GOP structure, which is described by two parameters GOP(N,M), where N defines the GOP length (i.e. the number of frames of each GOP) and the M-1 is the number of B frames between I-P or P-P frames, and taking under consideration the decoding inter-dependencies among the three frame types, then the impact of the packet loss ratio will be mathematically and deterministically formulated. In [28], it is presented the described theoretical mathematical model, which for given packet loss ratio and packet loss scheme provides the worst case theoretically expected decodable frame rate, without considering the decoding and error concealment capabilities of each decoder. The proposed model of the theoretically expected is summarized in the following equation

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

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

Where C_I C_P C_B are the mean number of packets that transport the data of each frame type, p is the packet loss rate, N_{GOP} is the total number of GOPs in the video flow, Ndec is the total number of decodable frames in the video flow, N_{dec-I} N_{dec-P} N_{dec-B} are the number of decodable frames in each type and $N_{total-I}$ $N_{total-P}$ $N_{total-B}$ are the total number of each type of frames.

The validity of this theoretical and mathematical framework has been examined in [28] by performing experiments using the ns-2 simulation platform for uniform packet loss distribution. As it has been deduced and shown on the cited paper, the dependence of the theoretically expected decodable frame rate and packet loss rate can be successfully described by an equation of the following form:

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

$$Q = C_1 \mathrm{Ln}(p) – C_2$$

Especially for the case of packet size equal to 1000 bytes and based on the performed simulation, the above equation is specialized to

$$Q = -0{,}3211\mathrm{Ln}(p) - 0{,}4094 \text{ with } R^2 = 0{,}9971$$

For the derivation of the above equation the random uniform model has been used, which provides the distributed losses with the mean loss rate (p) and corresponds to the worst case packet loss scenario, given that we have considered DT equal to 1.0. Regarding the AppQoS to the PQoS mapping, the application-layer Q metric can be extended to the service level, by expressing it in terms of duration percentage for error-free video transmission.

4. Conclusions

Digital video coding techniques have already prevailed in the upcoming broadcasting services and applications, enabling the provision of digital video content over various bandwidth-limited means and computationally low terminals. Video compression algorithms exploit the redundancy that a video signal contains in the spatial, temporal and frequency domain. Thus, by removing this redundancy in these three different domain types, it is achieved high compression of the data with cost the perceptual degradation of the content.

This chapter outlines the various PQoS evaluation methods and comments their efficiency. These methods can be mainly categorized into two major classes: The subjective and objective ones. The subjective test methods involve an audience of people, who watch a video sequence and evaluate its quality as perceived by them, under specific and controlled watching conditions. The objective methods successfully emulate the subjective quality assessment results, based on criteria and metrics that can be measured objectively. These objective methods are classified, according to the availability of the original video signal to Full Reference, Reduced Reference and No Reference.

Finally, the chapter discusses how the transmission errors and impairments of the transmission channel are mapped to the various QoS-related

Author’s Accepted Manuscript (AAM) Version: Chapter contribution in “*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*“, Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

layers of the broadcasting service. More specifically it is examined how the PQoS, AppQoS and NQoS layers are crossed related when transmission predicaments are present. In this context, the effect of the packet loss ratio on the theoretically expected ratio of decodable frames is discussed, describing how the interdependencies of the encoded frames create error propagation.

5. Acknowledgement

Part of the work in this chapter has been supported and was carried out within the framework of the Information Society Technologies (IST) Integrated Project ENTHRONE phase II/ FP6-38463.

6. References

1. Richardson I E G (2004) *H.264 and MPEG-4 Video Compression*, ed. Wiley.
2. Engelke U, Zepernick H-J (2006) *Perceptual Quality Measures for Image and Video Services.* Euro-NGI Workshop on Socio-Economic Aspects of Next Generation Internet, Lyngby, Denmark, October 9-10, 2006.
3. ITU (2000) *Methodology for the Subjective Assessent of the Quality of Television Pictures.* Recommendation ITU-R BT.500-10.
4. ITU (1999) *Subjective video quality assessment methods for multimedia applications.* Recommendation ITU-T P.910.
5. Arriba C (2007) *Subjective Video Quality Evaluation and Estimation for H.264 Codec and QVGA Resolution Sequences.* Institut fur Nachrichtentechnik und Hochfrequenztechnik Fakultat fur Elektrotechnik und Informationstechnik, Technischen Universitat Wien.
6. Silva E A, Panetta K, Agaian S (2007) *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 6579.
7. Wang Z., Sheikh H.R. and Bovik A C. (2003) *Objective video quality assessment*, in *The Handbook of Video Databases: Design and Applications*, B. Furht and O. Marqure, Editors, CRC Press. p. 1041-1078.
8. VQEG. (2000) *Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment*.
9. Wang Z, Bovik A C, Lu L (2002) *Why is image quality assessment so difficult?* in *IEEE International Conference on Acoustics, Speech, and Signal Processing*.

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

10. Wang Z, Lu L, Bovik A C (2004) *Video Quality Assessment Based on Structural Distortion Measurement.* Signal Processing: Image Communication, special issue on Objective video quality metrics 19(2): p. 121-132.
11. Wang Z, Bovik A C, Sheikh H R, Simoncelli E P, *Image quality assessment: From error visibility to structural similarity.* IEEE Trans. on Image Processing, 13(4): p. 1-14.
12. Watson A B (1998) *Toward a perceptual video quality metric.* Proceedings of SPIE Human Vision and Electronic Imaging, 3299: p. 19-147.
13. Lu L, Wang Z, Bovik A C and Kouloheris J (2002) *Full-reference video quality assessment considering structural distortion and no-reference quality evaluation of MPEG video*, in *IEEE International Conference on Multimedia*.
14. Gunawan I P and Ghanbari M (2003) *Reduced-reference picture quality estimation by using local harmonic amplitude information*. in *London Communications Symposium 2003.*
15. Ries M, Crespi C, Nemethova O, Rupp M (2007) *Content Based Video Quality Estimation for H.264/AVC Video Streaming.* Proc. Proceedings of IEEE Wireless and Communications & Networking Conference, Hong Kong.
16. Lu L *et al. (*2002) *Full-reference video quality assessment considering structural distortion and no-reference quality evaluation of MPEG video*. in *IEEE International Conference on Multimedia*.
17. Ries M, Nemethova O, Rupp M (2007) *Motion Based Reference-Free Quality Estimation for H.264/AVC Video Streaming.* International Symposium on Wireless Pervasive Computing , San Juan, Poerto Rico, 5-7 February.
18. Seeling P, Reisslein M and Kulapala B (2004) *Network Performance Evaluation Using Frame Size and Quality Traces of Single Layer and Two Layer Video: A Tutorial.* IEEE Communications Surveys & Tutorials, 2004. 6(3): p. 58-78.
19. Koumaras H, Kourtis A and Martakos D (2005) *Evaluation of Video Quality Based on Objectively Estimated Metric.* Journal of Communications and Networks **7**(3): p. 235-242.
20. Koumaras H, Pallis E, Xilouris G, Kourtis A, Martakos D, Lauterjung J (2004) *Pre-Encoding PQoS Assessment Method for Optimized Resource Utilization*. in *HET-NETs04* Ilkley, West Yorkshire, U.K.
21. Koumaras H, Kourtis A, Martakos D., Lauterjung J (2007) *Quantified PQoS Assessment Based on Fast Estimation of the Spatial and Temporal Activity Level.* Multimedia Tools and Applications 34(3): p. 355-374.
22. Koumaras H, Kourtis A (2007) *Video Quality Prediction Based on the Spatial and Temporal Classification of the Uncompressed Content.* The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Comm. (PIMRC), Athens, Greece, 3-7 Sep.
23. Koumaras H, Pliakas T, Kourtis A (2007) *A Novel Method for Pre-Encoding Video Quality Prediction.* IST Mobile Summit, Budapest, Hungary.

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)

24. Koumaras H, Gardikis G, Kourtis A (2003) *Objective Evaluation of the Perceived Quality of Video Content.* IST Mobile Summit 2003, Aveiro , Portugal.
25. Kanumuri S, Cosman P C, Reibman A R, Vaishampayan V A (2006) *Modeling Packet-Loss Visibility in MPEG-2 Video.* IEEE Transactions on Multimedia, 2006. 8(2): p. 341-355.
26. He Z, Xong H (2006) *Transmission Distortion Analysis for Real-Time Video Encoding and Streaming over Wireless Networks.* IEEE Transactions on Circuits and Systems for Video Technology **16**(9): p. 1051-1062.
27. Lee W, Srivastava J (2001) *An Algebraic QoS-Based Resource Allocation Model for Competitive Multimedia Applications.* International Journal of Multimedia Tools and Applications **13**(197-212).
28. Koumaras H, Kourtis A, Lin C-H, Shieh C-K (2007) *A Theoretical Framework for End-to-End Video Quality Prediction of MPEG-based Sequences.* The Third Inter. Conf. on Networking and Services - ICNS07, Athens, Greece.

Author's Accepted Manuscript (AAM) Version: Chapter contribution in "*Mobile Multimedia Broadcasting Multi-Standards: Application Handbook*", Editor Dr. Fa-Long Luo, Springer Editions, Chapter 14, pp.417-434, ISBN: 978-0-387-78262-1, October 2008 (Link)