Alexander TEI of Thessaloniki
School of Technological Applications
Department of Informatics

A study on Quality of Experience (QoE) for multimedia systems

Bachelor Thesis on Informatics

by
Vasilios Belias

Thesis Supervisor
Dr. Periklis Chatzimisios

Alexander Technological Educational Institute of Thessaloniki
Department of Informatics

2013

INDEXES

Abstract ....page 3
Chapter 1: Introduction...page 4
Chapter 2: Definition of QoE and correlation with QoS ....page 12
2.1 Definition of terms "Quality" and "Experience" ....page 12
2.2 Definition of Quality of Experience ....page 13
2.3 Factors influencing QoE ....page 16
2.4 Relation between Quality of Experience and Quality of Service ....page 20
2.5 Correlation between Quality of Experience and Quality of Service...page 22
2.6 QoS parameters related with end-user ....page 24
2.7 Correlation of Quality of Experience and Quality of Service...page 27
Chapter 3: Evaluation methods...page 32
3.1 Subjective Testing..page 32
3.2 Objective Methods..page 34
3.2.1 Media-layer models...page 35
3.2.2 Parametric packet-layer models...........................page 37
3.2.3 Parametric planning model..................................page 37
3.2.4 Bitstream-layer model...page 38
3.2.5 Hybrid model...page 39
3.3 Network planning models...page 40
3.4 Frameworks...page 40
3.4.1 Subjective Frameworks.......................................page 40
3.4.2 Objective Frameworks...page 44
Chapter 4: Standardization activities..page 48
4.1 Introduction to standardization activities.............................page 48
4.2 Standards for Video Quality...page 49
4.2.1 ITU-T P.910...page 49
4.2.2 ITU-T J.bitvqm...page 51
4.2.3 ITU-T P.NAMS ....page 51

4.2.4 ITU-T J.143..page 52
4.2.5 ITU-T J.140..page 53
4.2.6 ITU-T FG IPTV-C-0210..page 54
4.2.7 ITU-T FG IPTV-IL-0006..page 54
4.2.8 ITU-T J.149..page 55
4.2.9 ITU-T J.341..page 56
4.3 Standards for Audio Quality..page 56
4.3.1 ITU-R BS.1534-1..page 57
4.4 Standards for Speech Quality...page 58
4.4.1 ITU-T P.501..page 58
4.4.2 ITU-T P.564..page 59
4.4.3 ITU-T P.800..page 59
4.4.4 ITU-T G.107..page 61
4.4.5 ITU-T G.114..page 62
4.4.6 ITU-T P.851..page 63
4.4.7 ITU-T P.862..page 63
4.4.8 ITU-T P.800.1...page 64
4.5 Standards for Multimedia Quality.......................................page 65
4.5.1 ITU-T J.148..page 65
4.5.2 ITU-T G.OMVAS..page 66
4.5.3 ITU-T FG IPTV-C-0411..page 67
4.5.4 ITU-T G.1080...page 68
4.5.5 ITU-T FG IPTV-C-0354..page 68
4.5.6 ITU-T FG IPTV-C-0507..page 69
4.6 Synopsis...page 69
Chapter 5: Quality of Experience and applications...............................page 73
5.1 Application Areas for Quality of Experience.........................page 73
5.2 Cloud and Quality of Experience..page 75
5.3 IPTV and Quality of Experience...page 82
Chapter 6: Conclusions and future challenges....................................page 88
6.1 Conclusion...page 88
6.2 Future challenges...page 90
References...page 93

ABSTRACT

Quality of Service (QoS) was until recently the only tool in the hands of communication system researchers. But as we enter in a user-centric era where quality possesses a significant role, a new notion appeared, namely Quality of Experience (QoE). Conversely to Quality of Service, QoE is not concentrates in system-related characteristics, such as delay and jitter, but focuses in characteristics like behavior and perception. Quality of Experience is defined as "*an extension of the traditional QoS in the sense that QoE provides information regarding the delivered services from an end-user point of view*". The rest of this thesis is organized as following.

In the chapter 1 it is realized an introduction to the notion of Quality of Experience which is actually in a higher abstraction level compared to Quality of Service and therefore it could be considered as a perceptual pseudo-layer, which is positioned prior of the Application Layer of OSI Model.

In chapter 2 are presented some definitions for QoE stemming from different organizations, as well as, which factors affecting it. Moreover are referred the differences between Quality of Experience and Quality of Service alongside with the way these two are correlating.

In chapter 3 are being described QoE's assessment methods. These methods are: (i) subjective testing, (ii) objective methods and (iii) network planning models. Additionally, in this chapter are presenting frameworks which are using subjective testing and objective methods so as to assess Quality of Experience.

In chapter 4 are being presented the standardization activity concerning QoE. Specifically, the standards are categorized into the following four groups: (i) standards for video quality, (ii) standards for audio quality, (iii) standards for speech quality, (iv) standards for multimedia quality.

Finally, in chapter 5 there is a reference to the application areas of QoE, such as web and cloud, multimedia learning and sensory experience. Moreover, there is a specific reference to QoE in cloud and in IPTV.

CHAPTER 1

Introduction

Until lately, any research concerning communication systems was mainly conducted in the light of Quality of Service (QoS) [1] and consequently its objectives was the target area where international organizations, such as IETF and 3GPP, were focused [2]. A definition of QoS is given in [2] and is the following:

"*QoS is a measure of performance at the packet level from the network perspective and performance of other devices involved in the service.*"

Additionally, QoS includes a number of technologies, also known as QoS mechanisms, which are a useful tool in the hands of network administrators as with their assistance they are able to cope with the presence of any potential congestion that will be occurred on the application performance and moreover with those mechanisms it is possible to handle differently specific network traffics flows or a number of end-users by supplying another, probably better, level of service [2]. In layman terms, we can describe QoS as the capability to offer prioritization among different end-users, different applications and data flows. Furthermore, the parameters being used in order to guarantee a certain level of QoS are network-centric such as jitter, delay, bit error rate, bit rate, packet loss rate, etc and they are able to influence the perceived quality of a specific service that an end-user would experience, but nevertheless they cannot clearly quantify and subsequently measure this impact [1].

As the field of multimedia is keep developing continuously, a positive fact that it is being observed is that the quality of the upcoming products and services is being increased. Therefore, today when evaluating a product attention does not focus exclusively on the features which consist the product, but as well as if these features are addressed properly to the end-users and moreover what impact have on them [1]. In other words, normal users do not worry about the technology that it is being use by a product or a service, but their attention is concentrating mostly whether this product or service provides sufficient solutions to specific problems and what they would experience whilst utilize it [6]. This means that a new era is coming, an era of user-centric multimedia, in which quality would be a key factor and it is being expressed by a new notion, Quality of Experience (QoE). On contrary to QoS, which depends on system-related characteristics (e.g. jitter, delay), QoE considers characteristics such as perception, expectation, behavior, usability, needs, context, etc [1]. Therefore, the interest on the interaction between human (end-user) and computer has altered its attention from efficiency and effectiveness to factors that influence end-user's experience (e.g. engagement, enjoyment, etc) [4].

There is a plethora of definitions concerning QoE stemming from either from international organizations such as ITU and ETSI, or from other literature [4]. In [3] QoE is defined "*as the overall performance of a system from the users point of view*" and "*as the totality of QoS mechanisms, provided to ensure smooth transmission of audio and video over IP networks*". In [5] is described "*as the characteristics of the sensations, perceptions, and opinions of people as they interact with their environments. These characteristics can be pleasing and enjoyable, or displeasing and frustrating*" and "*as the user's perceived experience of what is being presented by the Application Layer, where the application layer acts as a user interface front-end that presents the overall result of the individual Quality of Services*". Finally, in [4] it is being defined as "*A measure of user performance based on both objective and subjective psychological measures of using an ICT (Information and communication technologies) service or product*".

From all the definitions mentioned above, it can be concluded that QoE possesses a higher abstraction level in comparison to QoS and subsequently it is possible to be reckoned as a perceptual pseudo-layer, added after the Application Layer of OSI Model. Based on the latter QoE can be also considered as an extension of QoS, because it supplies information concerning the delivered services as they perceived by an end-user [5].

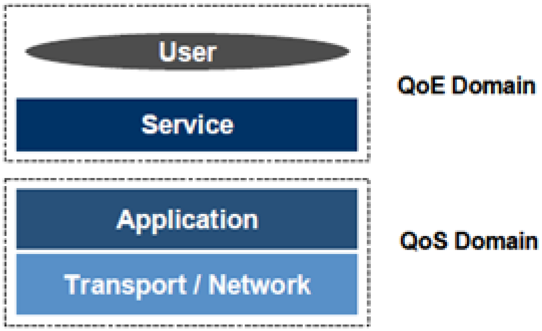

Figure 1.1 QoE/QoS layered model [5]

The layers presented in Figure 1, are the most significant from the QoS/QoE point of view. Those three levels [5] are described below:

1. ***Service Layer***: This layer corresponds to the perceptual pseudo-layer which is placed after the application layer and it is actually where the evaluation of QoE is taking place, as this is the layer which is exhibited to the end-user, either with subjective metrics which consider the impact that a specific service have to an end-user (e.g. MOS) or with objective metrics which try to approximate and predict end-users' opinion.
2. ***Application Layer***: In this layer parameters are correlated to a specific application (for example in case of video those parameters would be resolution, color, frame rate, video or audio codecs, etc) are being handled in so as to fulfill the desired level of QoE.
3. ***Transport/ Network Layer***: in this layer it takes place a management of parameters that are linked to transport and network performance such as jitter, packet loss and delay (actually they are the same parameters that affect QoS) in order to achieve decent QoE level.

Even though QoE is a relatively new concept in ICT, in other fields such as health and food sector it has been used a lot time ago as they were applied qualitative approaches so as to anticipate the desires of the consumers. Moreover, in ICT for a long period of time, it was considered that QoE could only be calculated subjectively [8], as experience itself is subjective and it depends on end-user's personal experiences, cultural background and from his social or economical status (e.g. if an interface is differentiated just by a few colors, it would have different impact to different people) [6]. Additionally, it is consumer's level of contention, concerning a specific service, which can potentially quantify or evaluate QoE [8]. Hence, in order to assess the QoE subjectively, experiments are conducted with the participation of human testers, under scientific rigor in order to produce trustworthy and valid results [1]. In those experiments, the participants are usually asked to evaluate the quality that they experienced using a service on a scale such as Mean Opinion Score (MOS) or inform about their ability to handle a service or to provide feedback about their level of satisfaction via survey techniques like questionnaires and interview or focus groups [4].

However, subjective testing has some drawbacks. It demands a lot of time in order to be conducted and has high cost, as well as, it cannot be applied in real-time applications. Consequently, this situation focused the attention of the research in the development of objective metrics and algorithms, which would be able to predict QoE as it would be by a human being, based on system characteristics that can be measured [1].

Nevertheless, so that to achieve the possible best reliability in QoE, it would be ideal to combine both subjective testing and objective metrics, namely we should correlate the subjective quality that an end-user experiences and the objective measurable quality [2]. For example, the efficiency and the effectiveness could be evaluated objectively alongside with the subjective user's satisfaction level [4].

QoE can be described by a variety of factors in a definition; nonetheless this does not create any problems if those factors are objective and can be easily assessed. However, it is difficult to accomplish such thing as QoE involves a lot of subjectivity and ergo puts various challenges to the multimedia community [6]. According to [9] (at which it is being attempting a report of those factors) an end-user's judgment, concerning the quality, for a system can be affected by "*pragmatic and hedonic*" aspects, which have to be assessed with the assistance of real or test end-user's, in order to anticipate the impact that they have on them. Those aspects are:

1. ***Interaction Quality***: It contains the input and the output perceived quality of the system. Input quality it is being defined as "*the perceived system understanding and input comfort*" and output quality as "*the perceived system understandability and form appropriateness*", whilst cooperativity contains "the distribution of initiative between the partners, the consideration of background knowledge and the ability for repair and clarification". Generally, interaction quality can be measured with questionnaire frameworks like, taking into account the following characteristics: smoothness, speed or pace, conciseness and naturalness of the interaction.
2. ***Efficiency-related aspects***: It contains the effectiveness, which depicts whether the precision and the completeness with which distinct end-users are able to accomplish certain targets in a given environment. It also contains intuitivity, which describes the level of efficacy that an end-user can actually reach when interacting with a technical system just by using knowledge unconsciously, as well as learnability, namely how fast and easy an end-user would be able to fully control the system. Finally it includes efficiency, which is the effort and resources that are demanded given a specific precision and completeness. For the measurement of efficiency-related aspects are being used questionnaires.
3. ***Usability***: A definition given by the ISO, describes usability as the "*extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use*". This aspect is further divided in two sub-aspects, the ease of use and the joy of use. The former is affected by the referred consequences of an interaction quality and the latter, beyond the quality of interaction, it is being influenced by hedonic aspects the system's "personality". Nonetheless, both those sub-aspects contribute to end-user's level of satisfaction. Finally, this aspect is also being measured with the assistance of questionnaires.
4. ***Aesthetics, system personality and appeal***: Firstly, aesthetics is being defined as the sensory experience that the system extracts and the level of personal targets that are being fulfilled by this experience.

Secondly, system personality is the extent to which the end-user comprehend the system characteristics that are stemming from the combination of agent factors and surface form. Finally, the appeal is a combination of the aesthetics of a product or a service, of its physical factors and of whether product's features are or not novel, surprising or fascinating.

5. ***Utility and usefulness***: If we desire to comprehend how useful is a system, we must make a comparison between the functional requirements that an end-user has and the functions that a specific system provides. On the other hand utility can be defined by the reply to the following question: "Is an end-user able to give a solution to his personal task with the assistance of the system?"
6. ***Acceptability***: It shows how easy is for an end-user to utilize the system. Moreover, acceptability can be used as an economic measure by linking the total of possible end-users to the quantity of the target group.

Furthermore, in [7] there is a reference to the people, who are labeled as "stakeholders", and have professional interest at employing data concerning QoE. These people can benefit from QoE as they can develop better products of services if they anticipate the end-users desires for a specific service (Table 1.1). They are categorized into three groups:

1. Service providers
2. Network operators
3. Manufacturers who develop equipment and software products

The reason for which QoE is important to stakeholders is because with its assistance are able to prevent products from being rejected. It observed that products was turned down from consumers, without being obvious to marketing departments the reason which was the low level of QoE. Those rejection would be probably predicted and avoided if the products were tested before being launched to the market by test users and the results of QoE where applied [7]. Additionally, it is also important for a service provider not to wait for an end-user to complain about a specific service and afterwards to conduct an evaluation of its quality, but on contrary it is vital to measure QoE perpetually and when it is necessary to ameliorate it [2]. This is underlined by the outcome of a research [7] which indicates that:

- a very significant percentage (82%) of end-users/customers churning, when they get irritated from the use of the product or a service
- an equally significant percentage (90%) do not protest before churning
- a disappointed customer/end-user will inform 13 more people concerning his bad experience from the use of a specific product

Technically oriented	***Customer oriented***	***Management***
Product Strategist	Marketing	CEO (Chief Enterprise Officer)
Strategic Service Developer	Service portfolio specialist	CTO (Chief Technical Officer)
System Integrator	Sales person	CMO (Chief Marketing Officer)
Project leader	Sales support	Chief of Strategy
SLA negotiators	Customer support	COO (Chief Operation Officer)
Technical researcher	Project leader	CRO (Chief Research Officer)
Interaction designer	Human factors researcher	CDO (Chief Development Officer)
Development engineer	Service host	Sales Director
Audio/Video codec engineer		

Table 1.1 Jobs that benefit from QoE use [4]

Furthermore, a service provider may notice that an end-user may experiences an excellent quality for a certain service, which has a specific level of QoS, while it is possible for a different level of QoS the same service to offer poor experience to the end-user. Nonetheless, it is not always guaranteed that when a service provider offers a high level of QoS that the end-user will subsequently experience high quality. It occurred that end-users where disappointed with an offered service, even though it had high QoS and consequently this led to low levels of QoE [10],[2].

Concluding, we can summarize some things that are known about QoE so far:

1. QoE and QoS are not similar [8]

2. QoS can be considered as a subset of QoE [2]
3. Both QoS and QoE can exist simultaneously, but this is not a guarantee that the one will lead to the other [2]
4. With the assistance of QoS only, it is not possible to create services/products that would consider end-user's experience [2]
5. The absolute target that QoS must have is to lead to high levels of QoE [8]
6. "*QoE is a function of behavioral variables that may be possible to measure quantitatively*" [2]
7. A number of those measures could be physiometrical, whilst some other could be psychometrical

CHAPTER 2

Definition of QoE and correlation with QoS

2.1 Definition of terms "Quality" and "Experience"

The term "Quality of Experience" it is being consisted of two words and each one of them it is also being considered as a separate term. The word "quality" and the word "experience". Definitions for those two terms are being given in [11] by Roto et.al (2011) and Jekosch (2005), so as to clarify them to the fullest and are following:

"***<u>Experience</u>:*** *An experience is an encounter of a human being with a system having defined beginning and end. Besides the temporal aspect, experience is influenced by the encounter's context i.e. its place and character. An experience can also include the experience of quality, but this is not a necessary prerequisite*."

"***<u>Quality</u>:*** *Is the outcome of a subjective evaluation process. It includes the transformation of the physical event into a perceptual event, the reflection about the perceptual event, the composition of the perceived features with some reference features, and the description of the quality event."*

It is crucial for both of the definitions analyzed above to consider them from end-user's point of view and hence, on contrary to *performance*, they cannot only be described by the accomplishment of a specific objective target or physical properties.

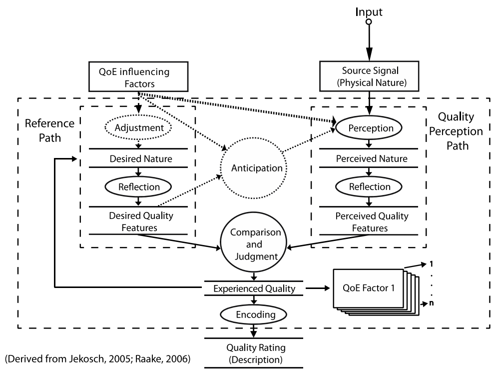

Figure 2.1 Quality formation process [11]

There are two distinct paths which participate to the quality formation process, as it is being depicted above, in Figure 2.1. The first one is the "reference path", which indicates the temporal and contextual nature of the quality process and moreover, acquires a memory of previous experienced qualities which are being pinpointed by the arrow headed from experienced quality towards to the reference path. The other one is the "quality perception path" which accepts as an input a physical event which may be triggered, for example, by a physical signal which our sensory organs perceive. The physical event, we just referred to, it is being processed via "low-level perceptual processes into a perceived nature of the signal" under the limitations of the reference path and afterwards the perceived nature it is being submitted a reflection process handled once more by the reference path, which "translates" these sensory features by cognitive processing.

At last, we can conclude referring that the combination of the wanted quality features which stemming from the "reference path" and understandable quality features coming from the "quality perception path" is forming the experienced quality on behalf of the comparison and judgment process. This experienced quality is delimited in space, time, and character and hence it can be labeled an event and since an event occurs inside the human user, consequently information concerning a specific event can only be gathered on a descriptive level coming from the side of user.

2.2 Definition of Quality of Experience

There are several definitions concerning Quality of Experience so in various internet sites as in a more formally expression from international organizations which are involved in information and communication technologies, such as ITU-T.

A first definition for QoE it is being given in [12] and is the following:

> **Quality of Experience (QoE)** measures total system performance using subjective and objective measures of customer satisfaction. It differs from Quality of Service (QoS), which assesses the performance of hardware and software services delivered by a vendor under the terms of a contract.

Moreover, it is being also mentioned that QoE model it is being utilized from both Information Technology (IT) and electronics industry for offering services and for business. Furthermore, as QoE lies on end-user/client experience, the level of user satisfaction it is being gathered from large user group polls. The outcome from the polls is influenced by the following factors:

- Efficiency
- Ease of use
- Reliability
- Customer loyalty
- Privacy
- Cost
- Security

Finally, there are referred and some environmental factors that possibly affect QoE such as hardware (e.g. wired or cordless devices), application criticality (e.g. texting versus audio/video) and lastly working environment (e.g. fixed or mobile).

Some other definition attempts for QoE are mentioned in [13] and are the four that mentioned below:

QoE as reloaded buzzword:

Quality of Experience has been defined as an extension of the traditional QoS in the sense that QoE provides information regarding the delivered services from an end-user point of view

QoE as usability metric:

Quality of Experience is how a user perceives the is how a user perceives the usability of a service when in a service when in use – how satisfied he/she is with a service in terms of, e.g., usability, accessibility, retainability and integrity

QoE as a hedonistic concept:

Quality of Experience describes the degree of delight of the user of a service, influenced by content network, device, application, user expectations and goals, and context of use

QoE as the ultimate answer to life, universe and everything:

Quality of Experience includes everything that really matters

Another definition it is being presented in [14] and it is following:

Quality of Experience (QoE or QoX) is a measure of the overall level of customer satisfaction with a vendor. QoE is related to but differs from Quality of Service, which embodies the notion that hardware and software characteristics can be measured, improved and perhaps guaranteed. In contrast, QoE expresses user satisfaction both objectively and subjectively. The QoE paradigm can be applied to any consumer-related business or service. It is often used in information technology (IT) and consumer electronics.

According to Wikipedia [15] QoE is defined as undermentioned:

Quality of experience (**QoE**), sometimes also known as **quality of user experience**, is a subjective measure of a customer's experiences with a service (web browsing, phone call, TV broadcast, call to a Call Center). Quality of Experience systems will try to measure metrics that customer will directly perceive as a quality parameter.

According to [16] QoE is defined as:

QoE: “The overall acceptability of an application or service, as perceived subjectively by the end user.” (ITU-T)

Note 1: Includes the complete end-to-end system effects

Note 2: May be influenced by user expectations and context

Another definition concerning QoE has been established at the Dagstuhl Seminar 09192 "From Quality of Service to Quality of Experience“(May 2009) and describes it as:

the degree of delight of the user of a service. In the context of communication services, it is influenced by content, network, device, application, user expectations and goals, and context of use. (cited after Möller, 2010)

Also in the same seminar is referred a definition for service, a word which is mentioned in the definition of QoE and follows:

Service: “*An event in which an entity takes the responsibility that something desirable happens on the behalf of another entity.” (cited after Möller, 2010)*

It should be mentioned that the definitions of Acceptability and QoE have to be distinguished in terms of the “characteristic of a service describing how readily a person will use the service”:

Acceptability: “Acceptability is the outcome of a decision which is partially based on the Quality of Experience.” (Dagstuhl Seminar 09192, May 2009, cited after Möller, 2010)

Additionally, QoE have to be also distinguished from Performance:

Performance: “The ability of a unit to provide the function it has been designed for.” (Möller, 2005)

Finally, in [11] it is being presented the working definition of QoE which has been created based on older definitions that have been referred previously in this chapter and describe QoE as undermentioned:

Quality of Experience (QoE) is the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user’s personality and current state. In the context of communication services, QoE is influenced by service, content, network, device, application, and context of use.

2.3 Factors influencing QoE

An influence factor in [11] it is being defined as:

Influence Factor: Any characteristic of a user, system, service, application, or context whose actual state or setting may have influence on the Quality of Experience for the user.

There could be a specific number of Influence Factors which may be named by end-users in terms of QoE features and it would be rather preferable not to consider them as distinct, as they may be connect to each other. Those factors can be classified into the following three categories.

The first one of those is the "Human Influence Factors (IF)" and it is being defined as undermentioned:

Human IF is any variant or invariant property or characteristic of a human user. The characteristic can describe the demographic and socio-economic background, the physical and mental constitution, or the user's emotional state.

Human Influence Factors (IF) are complex and moreover they are inextricably linked to each other. They may have the ability to affect perceptual process at two significant levels. The first one is the level of the early sensory, which is also known as low-level processing, and at this level properties linked to either the emotional or the mental or the physical constitution of the end-user may play crucial role. These characteristics can be either dispositional, such as end-user's auditory and visual acuity, age and gender, or variant and more dynamic, such as lower-order emotions, user's mood, personality traits, motivation and attention level. The second level is of higher-level cognitive processing, interpretation and judgment and at this level other human affecting factors are significant. Once again these properties can either be invariant or relatively stable, such as socio-economic situation, education background, attitudes and values, personality traits, or variant and more acute, such as expectations, needs, knowledge, previous experiences and emotions.

Second one in this categorization are the "System Influence Factors (IF)" and are described by the following definition:

System IFs refer to properties and characteristics that determine the technically produced quality of an application or service. They are related to media capture, coding, transmission, storage, rendering, and reproduction/display, as well as to the communication of information itself from content production to user.

System Influence Factors (IF) can be sub-categorized into the following four groups:

- Content-related System IFs
- Media-related System IFs
- Network-related System IFs
- Device-related System IFs

The first sub-category is addressed to the content type, e.g. specific temporal or spatial requirements, color depth, texture, 2D/3D. The second one, which concerns "Media-related Systems IFs", is referring to media configuration factors, such as encoding, resolution, sampling rate, frame rate, media synchronization. The "Network-related Systems IFs" sub-category is mentioned to data transmission over a network, e.g. bandwidth, delay, jitter, loss, error rate, throughput. The last one group, which includes the "Device-related System" influence factors, is referring to those devices or end systems which are involved to an end-to-end communication path, including system specifications (e.g. interoperability, personalization, security, privacy), device capabilities (e.g. display size, screen resolution, color depth, user interface capabilities, loudspeakers, headphones, luminance, audio loudness, computational power, memory, battery life-time), equipment specifications (e.g. type/complexity/usability, ergonomic aspects, mobility) and finally provider capabilities and specification (e.g. server performance and availability).

The last category of influence factors (IF) is the Context Influence Factors (IF) and are being defined undermentioned:

Context IFs are factors that embrace any situational property to describe the user’s environment in terms of physical, temporal, social, economic, task, and technical characteristics (Jumisko-Pyykkö et al., 2010; Jumisko-Pyykkö, 2011).

The Context Influence Factors (IF) is possible to occur in various levels of dynamism, such as static vs. dynamic, of patterns of occurrence, such as rhythmic vs. random, and of magnitude, such as micro vs. macro. Concerning the physical context, mentioned at the definition above, it describes the characteristics of space and location and also includes any movements within and transitions between locations. Furthermore, any temporal aspects of the experience, such as duration, time of day and frequency of usage, either of the service or the system, are part of the temporal context. The economic context, which is reffered to the description, includes the costs, the subscription type and the brandname of either the system or the service. Moreover, the two characteristics of task and social are parts of the perceived experience, which can be understandable either, focused or in a multitasking situation, such as task context, or, with the presence of either people or by alone, such as social context. At last, the two remaining characteristics, the technical and information context, present the connection between the system of interest and other similar systems and services including applications, such as availability of an app instead of the currently used browser-based solution of a service, networks, such as availability of other networks than the one currently used, devices, such as existing interconnectivity of devices over Bluetooth or NFC, or additional informational artifacts, such as additional use of pen and paper for better information assimilation from the service used.

In [17] there is another suggestion about factors that influence the QoE. According to it, a basic principle in order to succeed a certain level of Quality of Experience is the guarantee of a proper level of intrinsic characteristics engaged in a service delivery classified as Grade of Service (GoS), Quality of Resilience (QoR), and finally Quality of Service.

Brief definitions about these last three notions are mentioned below:

QoS is related to all phenomena occurring while the traffic is transported over the network.

GoS describes the process of connection setup, release and maintenance.

QoR encompasses all aspects of network survivability, dependability as well as service availability.

Quality of Experience is inextricably linked with the intrinsic network features and performance, despite the fact there is not a straight mapping between Quality of Experience and Quality of Service/ Quality of Resilience/ Grade of Service parameters and these relations are more comprehensible in context of particular applications. Nonetheless, there are efforts in progress so as to invent mathematical connection between intrinsic network parameters and Quality of Experience, commonly expressed quantitatively by the Mean Opinion Score (MOS) value. This last one is an extremely challenging task as QoE is also relied on many orthogonal factors such as the type of service, the end-user's terminal capabilities, the place in which the service is received, the pricing policy and the sociological and psychological aspects.

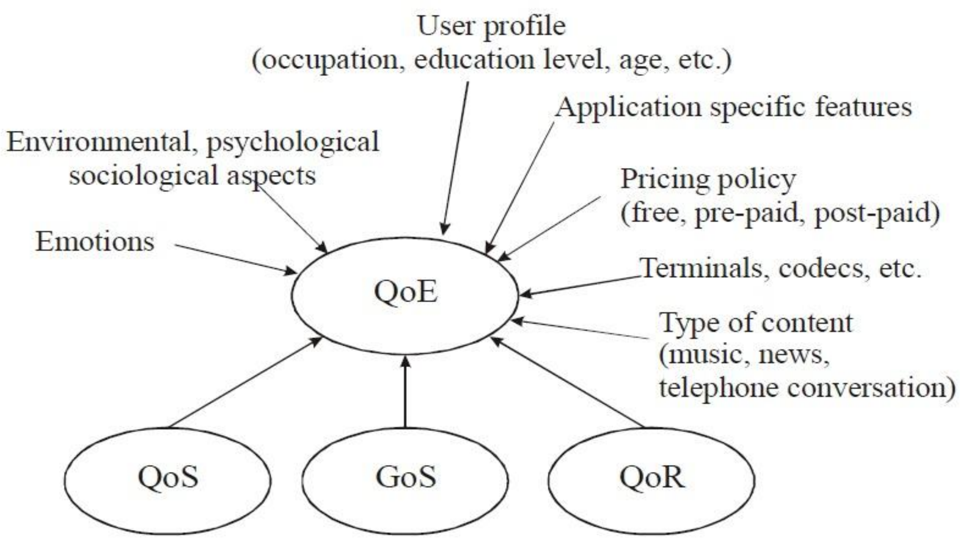

Figure 2.2 Factors influencing QoE [17]

2.4 Relation between Quality of Experience and Quality of Service

As it is being observed from the definitions which previously where mentioned in this chapter, there are significant differences between Quality of Experience and Quality of Service. In [11] there is a definition of Quality of Service defined by ITU (ITU-T Rec. E.800, 2008) and is following:

Quality of Service: “[The] Totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service.”

In this definition we can spot numerous clear differences with the notion of Quality of Experience and its defining characteristics. More specifically, we can observe that on contrary to Quality of Service which is concentrated in telecommunications services, Quality of Experience has a wider scope. Moreover, we can also mark that factors which are significant for Quality of Experience (i.e. the context of usage, ser characteristics, etc) and additionally the multi-dimensional nature of Quality of Experience are not properly addressed by Quality of Service in the definition above.

Nonetheless, a more significant distinction can be observed from the vast majority of research concerning the Quality of Service, which has a long tradition spreading over 20 years and is not being harmonized to the definition just mentioned from 'ITU-T Rec. E.800, 2008' nor to a previous definition referred at ' ITU-T Rec. E.800, 1994'. Alternatively, they are mainly focused to the notions of network performance and sometimes in other systems-level performance parameters.

Therefore, we can conclude that apart from the "official" definition of Quality of Service, there is a de-facto definition which distinct even more the last one from Quality of Experience as is more engaged with physical, quantitatively performance factors of networks and delivery platforms in general.

Hence, the difference between Quality of Experience and Quality of Service lies on numerous factors such as:

- ***Scope***: Quality of Service is mainly concentrated in telecommunication services, on contrary to Quality of Experience which in some cases does not even engage telecommunication (i.e. HD video played in home theater), as it covers a wider domain.

- ***Focus***: unlike to Quality of Service, which is engaged with performance aspects of physical systems, Quality of Experience concentrates most at end-user's evaluation of system performance, as colored by culture, context, the expectations that an end-user have from a system or a service and the level of satisfaction that he will have finally have from it, the psychological profiles and the socio-economic issues among other factors.
- ***Methods***: Quality of Experience demands a multi-disciplinary and multi-methodological approach for its comprehension, whilst Quality of Service demands a very technology-oriented approach and it depends on analytic approaches and simulative or empirical measurement.

After being pointed out the differences between QoS and QoE, it is crucial to recall that QoE is actually a large part of instances inextricably linked to QoS. The technical aspects of a system's performance, especially in case of multimedia systems, may have an important and in some cases defining impact on QoE dimensions.

2.5 Correlation between Quality of Experience and Quality of Service

The increasing growth of broadband IP networks has favored the development of the "Triple Play", which is the integration and convergence of voice services, such as Voice over IP (VoIP), of video services, such as Video on Demand (VoD) and data services. Consequently, as numerous services are offered in the integrated service environments, the notions of Quality of Experience and Quality of Service were appeared in the IP network.

Nowadays, end-user's satisfaction is one of the issues that interest the most the service and network providers. Nonetheless, the current methods which are used for the assessment of Quality of Experience are in most cases depended on scores from user and user survey, which are time consuming and costly, as well as they are too subjective. Therefore, this fact led to attempts in which they try to connect the objective network service conditions with the human perception of the Quality of Service.

Furthermore, as current Quality of Experience evaluations schemes assess Quality of Experience with the assistance of a part of Quality of Service measurement parameters, they are difficult to reflect of various services. The quality evaluation scheme that applies the QoS measurement parameter that already exists is able to supply the objective quality information in a network level. Nevertheless, this quality evaluation scheme is difficult to anticipate the level of quality that the end-user perceives from a service. However, attempts have been conducted in which, with the usage of Quality of Service information assessed in a network level it is being proposed a correlation model between Quality of Experience and Quality of Service.

An attempt it was conducted by Khirman and Henriksen, in which they tried to connect the human perception of quality of the service and the objective network service conditions. Their research has been widely examined for voice delivery and it is generally accepted that the relationship between the end-user understanding of quality and voice transmission conditions is away from linear as it is being mentioned in [18]. At their work, Khirman and Henriksen, dealt with how the level of satisfaction that an end - user experiences from an HTTP service, is influenced by two essential network Quality of Service parameters; the network delivery speed and the latency. Nevertheless, it is difficult to anticipate the level of user satisfaction for a service just only from the latency time and the bandwidth in an integrated network environment.

In another attempt in [19] the writers thought that the ubiquitous computing environment brings the method of evidence context connected to Quality of Experience. They examined the Quality of Experience assessment method in ubiquitous computing environment and suggest the enhanced Quality of Experience assessment parameter model. Moreover in [19] it is also being suggested a rough-set based algorithm in order to reduce context attributes and define the importance of each attribute, which has been tested on video streaming service. As a mass of evidence information connected to the experience of an end-user can be collected via the context-awareness computing, the calculation outcome of Quality of Experience assessment method can extremely anticipate the actual feeling of an end-user.

Nevertheless, this method has to be optimized in order to correspond to the growth of ubiquitous computing.

2.6 QoS parameters related with end-user

In this chapter are referred the Quality of Service quality parameters which can be taken into account for the assessment of Quality of Experience. Probably, the most essential of those parameters is the "Transfer Capacity", which plays a crucial role in the influence on the performance realized by the end-users. The majority of applications addressed to an end-user have a minimum in capacity requirements, which have to be taken into consideration when entering into service agreements and moreover any potential lost bits or octets should be removed from the total sent data, so as to calculate networks capacity.

Furthermore, it can be presumed that there is an agreement between the end-user and the network provider at the maximum access capacity that would be offered to one or more packet flows within a certain Quality of Service class, omitted the "Unspecified" class presented in Table 2.1. A packet flow definition is following:

A **packet flow** is the traffic associated with a given connection or connectionless stream having the same source host (SRC), destination host (DST), class of service, and session identification. Other documents may use the terms microflow or subflow when referring to traffic streams with this degree of classification.

Network Performance Parameter	QoS Classes					
	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5
IPTD	100 ms	400 ms	100 ms	400 ms	1 s	U
IPDV	50ms	50ms	U	U	U	U
IPLR	1x 10^{-3}	1x 10^{-3}	1x 10^{-3}	1x 10^{-3}	1x 10^{-3}	U
IPER	1x 10^{-4}					U

Table 2.1 IP network QoS class definitions and network performance objectives [20]

Firstly, both end-user and service provider are able to use whatever capacity specification they believe to be suitable for the case, as long as they permit both user verification and network provider enforcement (i.e. it may be enough to specify the peak bit rate on an access link, having considered the lower layer overhead too). Namely, the network provider accepts to transfer packet at a predefined capacity according to a specific Quality of Service class.

Nonetheless, there could be superfluous packet submission which could affect either the capacity agreement or the negotiated traffic contract and hence the performance objectives could not be applicable. So, in case that a superfluous is detected, the network has the possibility to reject so many packets as the total number of the excess packets. If a packet is rejected, then it should be eliminated from the population of interest, which is consisting from assessed packets with the assistance of network performance parameters. More specifically, rejected packets should not be considered as lost packets whilst evaluating the network's IPLR performance, as a rejected packet might be transmitted again and thus it would be treated as a new packet in the evaluation of the network performance.

Finally, every network's Quality of Service class produces a certain combination of limitations on the performance values (Table 2.2). A specific network Quality of Service class can be addressed under specific circumstances, but it is not obligatory the usage of any particular network Quality of Service class in any specific context.

Another Quality of Service quality parameter that influences Quality of Experience is the "Delay". It can be described in various ways, such as the time which is demanded in order to initialize a specific service from the primary user request and the time consumed until to accept certain information, since the service is established. Delay has an immediate impact on end-user's contentment depending on application, whilst the term "Delay" refers to any delay occurred either at the network or at terminal or in any servers. Moreover, we can remark that from an end-user perspective, delay also consider the potential influence of any other network parameters (i.e. throughput).

QoS Class	Applications (ex)	Network techniques
0	Real-time, jitter sensitive, high interaction (VoIP, VTC)	Constrained routing and distance
1	Real-time, jitter sensitive, interactive (VoIP, VTC)	Less constrained routing and distances
2	Transaction data, Highly interactive (Signaling)	Constrained routing and distance
3	Transaction data, interactive	Less constrained routing and distances
4	Low cost only (short transactions, bulk data, video streaming)	Any route/path
5	Traditional applications of default IP networks	Any route/path

Table 2.2 Guidance for IP QoS classes [20]

"Delay Variation" is the next quality parameter that affects Quality of Experience. It is mostly considered as a performance parameter as it is extremely important at the transport layer in packetized data systems because of the inherent variability that the individuals packets have in arrival time. Nevertheless, in case of services which have low levels of tolerance in delay variation, actions are taken in order to either remove the last one or at least to decrease it importantly with the assistance of buffering and hence to successfully eliminate the "delay variation" as it is being comprehended in an end-user level.

The final, Quality of Service quality parameter that affect Quality of Experience is the "Information Loss". This one plays a crucial role to the information that will at last exhibited to the end-user regardless if it is image, video, voice or data. "Information Loss" is not only caused by the bit errors or packet loss during the transmission but it can be caused as well by any possible degradation created by media coding in order to achieve more efficient transmission, such as the usage of low bit-rate speech codecs for voice.

2.7 Correlation of Quality of Experience and Quality of Service

The already existing Quality of Service/Quality of Experience correlation of the service can be appeared in many different ways into an integrated network environment from Quality of Service parameters considered for a type of service and Quality of Experience. In [20] a solution is suggested which is the correlation model approach method which has the ability to reflect all those elements that affect the Quality of Experience.

The connection between the Quality of Experience and Quality of Service is not easy to be mapped with the offered service within a converged network. Each Quality of Experience class so as to be satisfied it demands a different Quality of Service, depending on the traffic characteristics of each service. Furthermore, the Quality of Experience affecting element at the terminal layer, which is necessary so as to assess the last one, has to reflect. Moreover, the normalized Quality of Service score of the service which can be offered is able to be acquired through the following equation from the integrated network.

$$QoS = F(D, J, L, E, B, S) \quad (1)$$

So, with the usage of quality indexes of a network-level, the score for Quality of Service can be learned. It realized with application in order to exist delay, jitter, loss rate, error rate, bandwidth, the signal success rate and other, which are objective quality parameters and the weighted value of those parameters is not the same from the used service.

As an example, a service which is sensitive to delay time conveys the weighted value 10 to a delay time of 100ms (i.e. VoIP, video conference service), on contrary when a service is tolerating delay time (i.e. Video On Demand) it is being given the weighted value of 5. Moreover, is useless to provide bandwidth over the necessary if VoIP and video conference service can correspond with the minimum guaranteed bandwidth. Conversely, there are services in which the perceived quality from an end-user is better as the supplied bandwidth is augmented.

$$QoE(QoS) = K\left\{\frac{(e^{QoS-\alpha} + e^{-QoS+\alpha})}{(e^{QoS-\alpha} + e^{-QoS+\alpha} + \beta)} + 1\right\} \quad (2)$$

In this last equation (2), the involved constant and variables are defined by elements that affecting Quality of Experience considering the environment and the type of using service, the terminals position, the used codecs, etc. Additionally, the mentioned Quality of Service of this equation is actually the normalized score estimated from the equation (1). Furthermore, the Quality of Experience class calculated through the Quality of Service quality parameters of a network-level is matching as in the current Mean Opinion Score (MOS) grade, which has 5 classes. Moreover, the "a" which ought to be given as the Quality of Service class constant so as to content the demanded Quality of Experience. Continuously, "β" is defined based on the class of service as the grade of service constant. Lastly, "K" is the scale constant which indicates the level of the end-user contentment from the usage of a specific service.

According to [20], a classification can be realized for the offered services of an integrated network and is the following:

1. The Guaranteed Service
2. The Best Effort Service
3. The Premium Service

To the first one of them, the "Guaranteed Service", it is being given top priority. Hence the bandwidth guarantee, as well as, data priority processing, etc are indispensable according to the service priority so as to reassure the real-time multimedia service, such as VoIP, IPTV, Video Conferencing, etc.

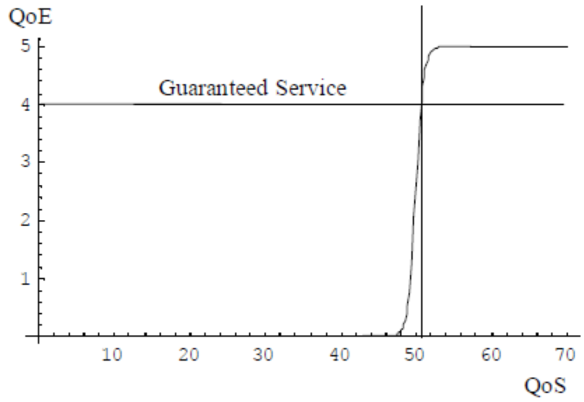

Figure 2.3 The QoS-QoE correlation model of the Guaranteed service [20]

The services that are most likely to be guaranteed are real-time voice or video services (i.e. VoIP, Video Phone, Video Conference, etc). In order to an end-user has a good experience for one of the services just referred, stable bandwidth is taken for granted and also must to be guaranteed that the connection delay time will be to its minimum. Furthermore, the quality parameters demanded for the quality evaluation of those services are the delay, jitter, packet loss, etc. Additionally, when these services take into account for the quality measurement, the Quality of Experience, the call connection yield, the usage of the service duration, as well as, either the audio or the image MOS value are necessary.

So, in the case of "Guaranteed Service" the Quality of Service/Quality of Experience correlation model can be introduced by the on/off model. Namely, when the demanded Quality of Service score for the guarantee of a certain level if Quality of Experience is not fulfilled, the "Guaranteed Service" is impossible to be offered. For example, as it can be understandable from the Figure 2.3, the Quality of Service score ought to at least 50 so as to reassure the Quality of Experience for the specific model. Moreover, as the "β" value of this model is small in comparison with other models leads us to the conclusion that the minimum guaranteed class of the Quality of Service which has to be provided is much more sensitive.

The next one is the "Best Effort Service" which incorporates the already existing internet services with the quality level equal to it and it is the more delicate of the Quality of Service/Quality of Experience correlation models. This signifies that even if the Quality of Service parameters will be ameliorated, there is an upper high in the perceived quality that a user can comprehend.

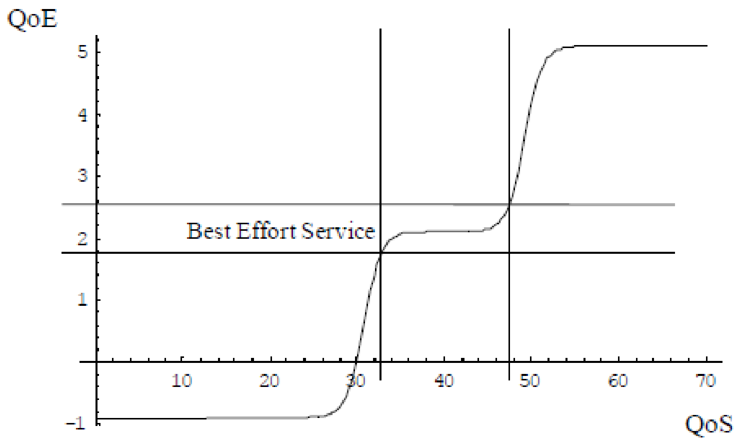

Figure 2.4 The QoS-QoE correlation model of the Best Effort service [20]

Finally, the last one in this classification is the "Premium Service", which is more tolerant to the real-timeness than the "Guaranteed Service". This category is mostly focused on services like IPTV, VoD, movies streaming service, interactive gaming and so on. Furthermore, concerning to the traffic characteristics that those services demand, we have to guarantee the real-time streaming service, as well as, the multicasting service and the minimum bandwidth. Additionally, the main quality parameters of "Premium Service" are same with those of the "Guaranteed Service", but nevertheless the first one is less sensitive to delay than the last one.

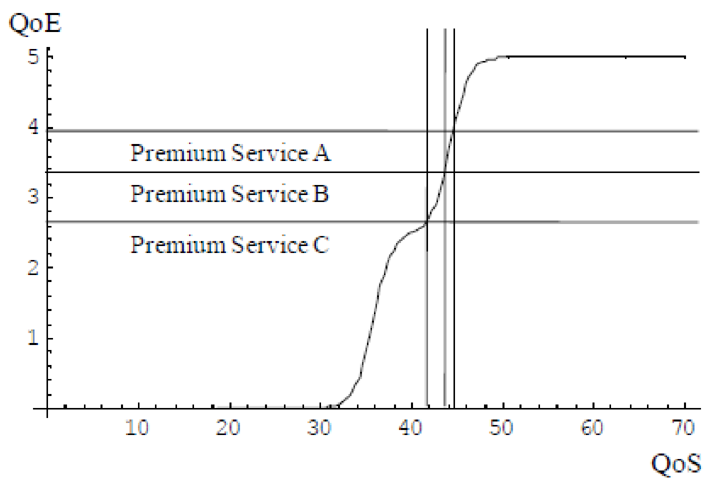

Figure 2.5 The QoS-QoE correlation model of the Premium service [20]

In case of this category, the Quality of Service/Quality of Experience correlation model has the characteristic that it is linear, and in fact more linear than the "Guaranteed Service and this can be adjusted via the grade of service constant "B". Concluding, in order to choose the suitable Quality of Experience class constant, an agreement can be realized between the network and the service providers so as to organize the different "Premium Service" Quality of Service/Quality of Experience correlation models.

CHAPTER 3

Evaluation methods

There are three categories in which the Quality of Experience's evaluation methods are classified. Specifically, those are the subjective testing, the objective modeling and the network planning.

3.1 Subjective Testing

On contrary to Quality of Service, which evaluates mostly, the technical characteristics that influence the service performance, Quality of Experience have to concern measures from users. Therefore, subjective QoE does not limited in media's quality that an end-user perceives, but it also contains measures like the level of end-user satisfaction and the usability [21]. Based on that, the QoE evaluation method of subjecting testing was developed.

In subjective tests, it is required the participation of human testers. Specifically, the concept is to gather the test participants in a room in order to assess a service under real conditions, normally by completing questionnaires. The combination of the answers and the questions can provide us with a variety of information concerning the testing service. Additionally, the most known way of evaluating the QoE in subjective tests is the Mean Opinion Score (MOS). MOS is used especially for video assessment and it is an absolute metric, namely it has no negative values in its scale (this is occurred in comparative metrics (Table 2)), but has a five point scale starting from 5 for "Excellent" to 1 for "Bad" (Table 1) [23], [24]. Moreover, the parameters of such an experiment like the selection of the test participants, as well as, the sample size of the testing service (e.g. video, audio, etc), the environment, etc, is essentially to be configured correctly. Guidelines for an appropriate preparation are available in standardization activities such as in ITU-R Recommendation BT.500-11, which presents a subjective evaluation method concerning video quality [22].

The way of conducting a subjective test which described above, it is labeled according to [25] "formal evaluation". Moreover, another type of evaluation it is being referred, labeled "informal evaluation". In this type of evaluation method the assessment of video quality it is realized by service provider craftsperson on spot, as well as, technical experts, also known as "golden eyes", either during the assignment or in video system head end. The advantage of this evaluation method is that the experts, who are extremely experienced, are aware of what they are seeking. On contrary, the negative side is that those experts are not always available and additionally they can possibly fail in anticipating the preferences of the end-users.

Generally, the most significant advantage of subjective testing is that it is the unique way to anticipate the psychological and sociological impact on QoE. Furthermore, it is probably the most reliable method for the evaluation of QoE, especially for the video service assessment [20].

Nevertheless, subjective testing has and its drawbacks. The most important of those is that they are time consuming and demand a specifically designed environment, which subsequently render them costly. Additionally, it is a very demanding task so the creation, as the execution of a subjective tests which will be able to supply statistically interpretable and recurring results [25]. Furthermore, another negative concerning the subjective testing is that for real-time environments are neither suitable nor scalable [26].

Examples of approaches appropriate for subjective testing is the Single Stimulus Continuous Quality Evaluation (SSCQE), in which the test participants are able to dynamically evaluate the quality of a randomly video sequence, which has long duration, with the assistance of a slider machine having an associated quality scale and also are the following: *" Double Stimulus Impairment Scale (DSIS), Double Stimulus Continuous Quality Scale (DSCQS), Single Stimulus Continuous Quality Evaluation (SSCQE), Simultaneous Double Stimulus for Continuous Evaluation (SDSCE) and Stimulus Comparison Adjectival Categorical Judgment (SCACJ) (Bocca-Rodríguez et al., 2007). "* [26]

Absolute metrics		
MOS	*Quality*	*Impairment*
5	Excellent	Imperceptible
4	Good	Perceptible
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Table 3.1 Mean Option Score [26]

Comparative metrics	
Score	*Description*
3	Much Better
2	Better
1	Slightly Better
0	About the same
-1	Slightly Worse
-2	Worse
-3	Much Worse

Table 3.2 Comparative metrics[26]

3.2 Objective Methods

As mentioned before, subjective methods are costly and time consuming. In order to overcome this obstacle, efforts have been made so as to create alternative options [27]. One alternative is the objective methods which have as purpose, according to [22], to eliminate the drawbacks of subjective testing. Even though, according to [25], objective methods are not totally substitute the subjective testing, but they rather function as complementary. As it is referred, an ideal combination of objective and subjective methods could lead as to collect as good results as it gets, concerning Quality of Experience.

Those methods are able to provide a QoE assessment for a specific service either through the measurement of a number of parameters or by examining the signal with quality indicators placing at the output of the transmission channel, achieving in that way, for example, to evaluate the signal as it would be actually judged by an end-user, after the impact it would have to him. Furthermore, in objective methods various features of the signal are being analyzed, e.g. if we measure the quality in video streaming, the result of the evaluation will depend from the existence of various degradations in video's image, namely whether there are frame skips and freezes or not , whether contrast and brightness are in a satisfying level, etc [22].

Additionally, objective methods can be categorized into three groups based on whether they use the original signal. Those categories are:

- *Full Reference (FR):* here, a comparison it is being conducted between the source signal (or otherwise called reference signal) and the received signal (or otherwise called distorted signal) and hence the video quality it is being specified objectively. Examples of FR are the PEVQ, which it is being described in ITU-T's Recommendation J.247, the PESQ, which it is being described in ITU-T's Recommendation P.862 and the SSIM (Structural Similarity Index), which is used in video transmission [22],[25].
- *Reduced Reference (RR):* in this approach, the quality it is being determined objectively by partial information about the reference signal and full information concerning the received signal [25].
- *No Reference (NR):* In many cases the reference signal might be unavailable. So, the NR methods give a solution to this problem by analyzing the received video signal only, so as to evaluate the quality objectively [22], [25].

Generally, objective quality evaluation methods can be into the following five types:

1. Media-layer models
2. Parametric packet-layer models
3. Parametric planning models
4. Bitstream layer models
5. Hybrid models

3.2.1 Media-layer models

Media-layer models utilize as input the original media signal, whether it is a video or audio and they might as well consider channel characteristics and codec compression. In order to calculate QoE they are using complicated "*perceptually-based psycho-physical models*" and they achieve that either by comparing the distorted signal to the reference signal, namely with the assistance of full-reference (FR) (Figure 3.1) and reduced-reference (RR)

(Figure 3.2) methods, or by studying only the received signal, namely with the assistance of no-reference (NR) (Figure 3.3) method [27].

Furthermore, it can be used in the assessment of unknown systems, for example in codec comparison/optimization, as it does not demand in advance knowledge of the testing system. On contrary, in cannot be applied in cases where media signals are not available, e.g. it is not possible to acquire media signals at a network's mid-point , even though could be possible to decrypt the payload of a packet [28].

Additionally, media-layer models in order to anticipate video quality as an end-user would do, they use knowledge of the human visual system, which can be described in two phases. The first one concerns psychophysical models which stand for primal and low-level optical information treatment, for example "*spatial and temporal frequency response*". The other one concerns cognitive models which are responsible for high-level functions such as "*memory biases and judgment operations*" [28].

Finally, cases where this model it is being applied are being presented in ITU-T's Recommendations J.144 (video) and P.862.1 (speech) [28].

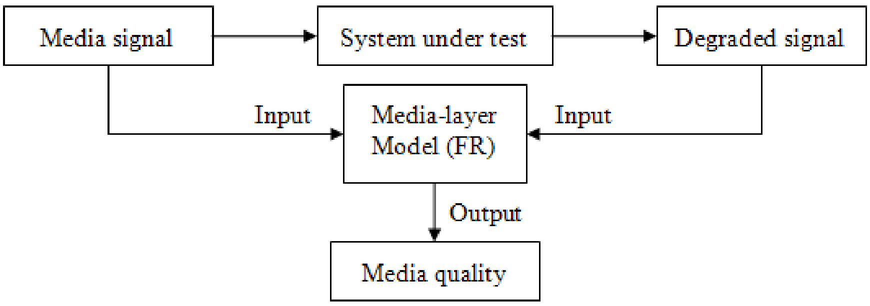

Figure 3.1 Media-layer model for full-reference method [27]

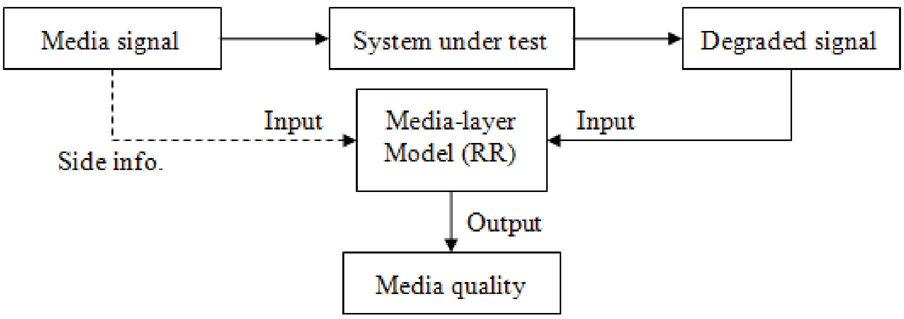

Figure 3.2 Media-layer model for reduced-reference method [27]

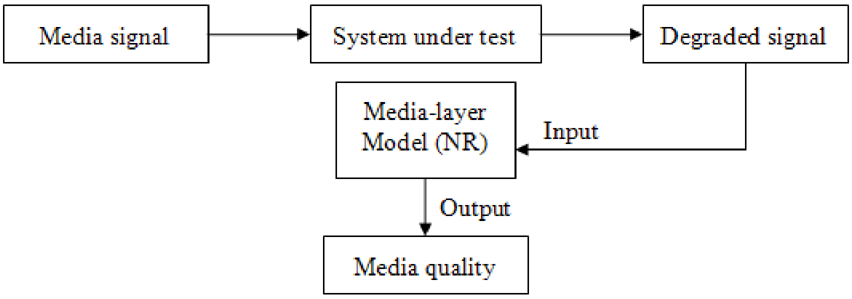

Figure 3.3 Media-layer model for no-reference method [27]

3.2.2 Parametric packet-layer models

Parametric packet-layer models are able to predict QoE exclusively from packet-header information and they have significantly light measurement of computational efficiency. As they skip the payload information, they are not able to assess the QoE in the media content. The main purpose of those models is to be used as probes either at the end-points or at the mid-points of a network. Finally, in ITU-T's Recommendation P.564 it is being presented the framework and the performance demands for this type of models [27], [28].

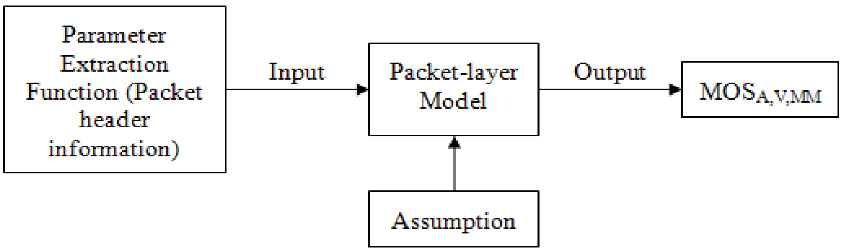

Figure 3.4 G.RQAM - Packet-based Model [27]

3.2.3 Parametric planning model

Parametric planning models are accepting as input quality planning parameters so for networks as for terminals. Conversely to the media-layer models, they do not demand in advance knowledge of the testing system [27], [28]. According to [28] a remarkable example of a parametric planning model is the E-model, which it is being presented in ITU-T's Recommendation G.107. E-model was broadly used in public switch telephone network (PSTN) and for voice over IP (VoIP) services as network planning tool. Finally, another ITU's standard that describes a model of this category is ITU-T Recommendation G.1070 [28].

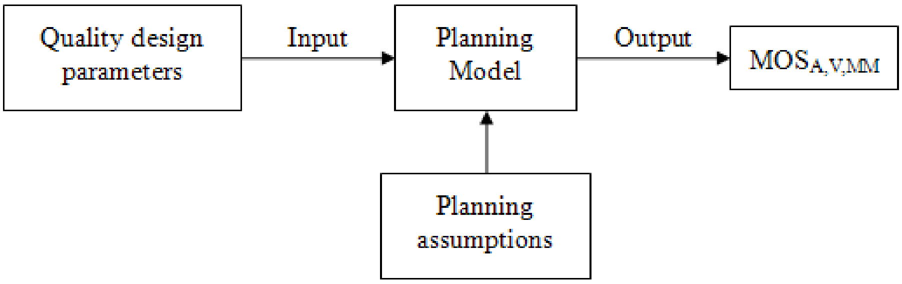

***Figure 3.5* G.RQAM - Planning Model [27]**

3.2.4 Bitstream-layer model

A bitstream-layer model is a relatively new concept and it is actually stands between the media-layer models and the parametric packet layer models, namely it is a combination of those two models [27], [28].

In case of "*in-service non-intrusive*" evaluation of QoE, parametric packet-layer models are efficient as they provide light computational load. But as mentioned before, they not examine the payload information and subsequently it is extremely difficult to anticipate the impact that video and audio have on the quality. So, the calculated quality for a parametric packet-layer model it is actually an average over a part of the content and this is inappropriate in case where someone has to observe the QoE of individual end-users. This issue can be overcome with the assistance of media-layer models as they are able to examine the characteristics of both audio and video content. Nevertheless, there is still an obstacle to surmount. It is difficult to acquire the media signal in QoE monitoring scenarios [28].

A solution to the problem described above, is to take advantage of the coded bitstream information in order to study the characteristics of source content. An example of this is given in [28] and is the following: " *DCT coefficients in MPEG-type coding tell us about the spatial complexity of video scenes, which affects the coding performance at a given bit rate and the robustness against packet loss*".

In brief, bitstream-layer model applies the encoded bitstream information in combination with the packet-layer information, so as to consider the content-dependent quality assessment characteristics by using light computational load [28].

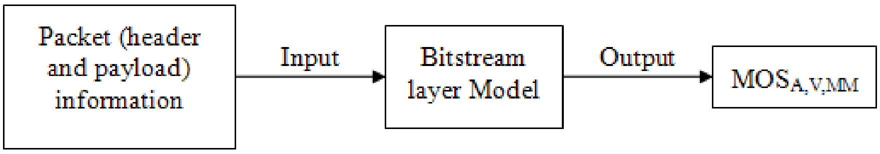

Figure 3.6 G.RQAM - Bitstream-layer Model [27]

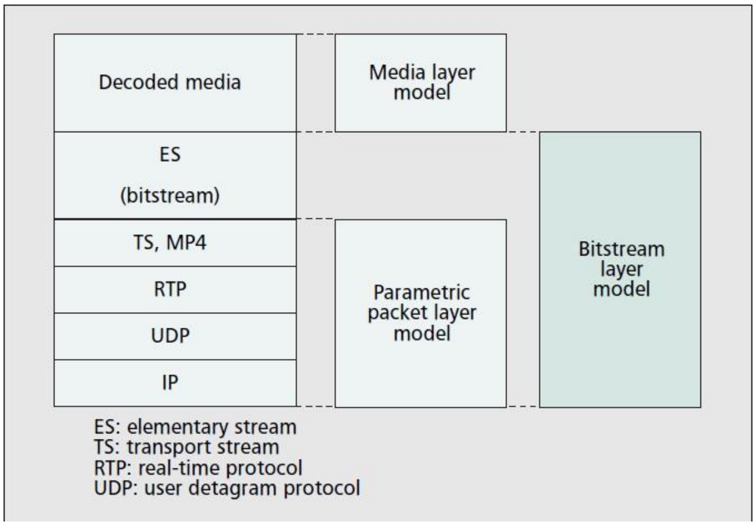

Figure 3.7 A comparison between bitstream model and media-layer, as well as, parametric packet-layer model [28]

3.2.5 Hybrid model

As it can be easily comprehended by its name, hybrid models are a combination of all models referred so far. An advantage that is has is that it can exploit the best possible information in order to predict QoE, e.g. utilize a media signal and bitstream information, having as purpose to gather information so from media-layer as from bitstream-layer models [27],[28].

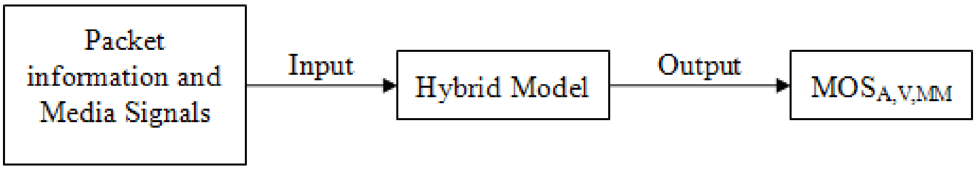

Figure 3.8 G.RQAM - Example of Hybrid Model [27]

Finally, concerning generally the objective methods, a significant advantage that they have is that they are quicker and cheaper than the subjective testing. On contrary, the main drawback that they have is that they may provide us with unreliable results. For example, it is widely known that evaluation methods such as Peak Signal-To-Noise Ratio (PSNR) may give results that have no relevance with human QoE assessment. Hence, an ideal solution to this problem would be to verify each new objective method through subjective testing, namely to check if they would provide similar results.

3.3 Network planning models

Network planning models does not require human testers in order to give a prediction of QoE. This is the reason for which, in some cases, they believed to be a subcategory of objective methods. A fact that separates them is that they are able to calculate QoE using a method which maps QoS measurable intrinsic quality parameters to QoE metric, instead of demanding to study the original signal. A typical example of those methods is the E-model [29]. Finally, as a subcategory of objective methods has the same positives and negatives that the last one have [22].

3.4 FRAMEWORKS

In this subchapter are being presented frameworks which are using the methods mentioned previously. Specifically, they are described frameworks that are use subjective evaluation, as well as, objective assessment.

3.4.1 SUBJECTIVE FRAMEWORKS

A framework that utilizes subjective testing it is being presented in [30]. Particularly, it is presenting a full length movie quality evaluation methodology which suggests that test participants should watch video sequences (it is proposed the use of full length DVD movies) in the environment that they normally watch their television programs. As the purpose of this framework is to examine IPTV and VoD the quality of the testing material, that is provided to test participants, must be of equivalent quality to the IPTV.

The DVDs that it is being provided to the end-users have some parts that are being on purpose damaged, so that would be feasible to be anticipated the impact that the audiovisual degradations have to the test participants. Subsequently, this methodology is suitable for testing the quality only in audio, or only in video or in both of them. Additionally, it mostly concentrates at real-life evaluation of QoE and hence it is preferred that the test participants have no knowledge concerning the parts of the movie where degradation exists. Immediately after they watching the movie they give their feedback concerning the quality that they experienced by completing questionnaires which are provided to them alongside with the DVD. Moreover, they are prompted not to read them before firstly watch the entire movie. Finally, some of the questions contained in the questionnaires are the following:

- "Did *you perceive any visual artifacts during playback? If yes, how many? Which types: motion jerkiness, blockiness, green blocks, other*"
- " *Indicate on a scale from 1 to 5 the annoyance of the impairment (1=not annoying at all, 5=very annoying)* " [31]
- "Which *types of impairments are the most annoying? (Motion jerkiness, blockiness, green blocks, other)?* "
- "Describe *the scenes or the locations where the degradations occurred. (e.g., the scene with the fire place at the beginning of the movie, when they are talking in close-up at the end of the movie)* "

With the assistance of this framework, two experiments were conducted in order to evaluate QoE under different circumstances.

In first case the DVDs that where provided to the test participants had degradation only in video quality, whilst the audio track stayed intact. Specifically, what was studying in this scenario was the impact that would have on the perceived quality by an end-user, the packet loss and the frame freezing. Generally, the purpose of performing this experiment was to determine whether visual impairments are equally observed during real life evaluation of QoE and moreover to also to point what type of impairments influence the most the viewing experience.

In the second one, it is being examined the impact that H.264 scalable video coding (SVC) has on the quality that an end-user perceive. Precisely, the reason for which this experiment has been conducted was to comprehend, as much as possible, the preferred scalability solution concerning the video downscaling. The video that was used for this experiment was a full-length movie, which has duration of 130 minutes and it was processed so as to have six degradations equally distributed, but not included in the first quarter of the film or in the last fifteen minutes of it. Additionally, the duration of each degradation would not exceed the eight seconds. Furthermore, the number of test participants was 38 and no one of them was participated in the first real-life QoE experiment which has as purpose to measure the affect that packet loss, as well as, frame freezing have on the quality that an end-user experiences. Moreover, they also had no idea that the movie would contain any impairment. Finally, after they watched the movie they had to fill a questionnaire which was the basis of a face-to-face interview they has the next day.

Another framework that uses subjective testing it is being referred in [32]. The target of this framework is to determine what impact has on the end-users a transition period during channel zapping in IPTV. In order to fulfill this goal the environment in which the test will be conducted must be similar to a typical TV environment. Furthermore, the intention of having a wide range of content for the needs of the test, was satisfied by the existence of six sequences (Table 3.3), whilst their resolution concerning the main stream was 640x360, while for the tune-in stream was a quarter of that. Moreover, both sequences where encoded and decoded with the assistance of H.264/AVC and especially the main stream was encoded in high quality, so that it would not have any potential impairments that could possible distract test participant's attention from the channel zapping.

Channel	***Description***
News	head and shoulder news speaker, mixed with short reports
Movie 1	close-up scenes mixed within moderate action scenes
Report	camera shots in town, with inline interviews
Movie 2	high motion action scenes
Football	scene from football game with some fast motion
Weather	weather forecast map with presenter, ticker with temperature numbers

Table 3.3 SEQUENCES USED IN THE TEST [32]

As mentioned, the attention of the test participants should be on channel zapping, namely at the tune-in process, which it is being described in three parts. Firstly, is the "*initial start-up delay*" at which a black screen is shown until the first image is exhibited. Next is the *"duration of the transition period",* which lasts maximum four seconds. The final one is "the level of tune-in quality Q during the transition period", which states that for good quality during transition the stream should be tune-in at a quantizer of QP26, whilst for not so good quality at QP30. Those three parts are being combined producing 18 unique scenarios (Table 4), which once the test starts are being displayed automatically on the screen, each on for 45 seconds. Then, the test participants have 10 seconds to evaluate.

Section	Scenario	Black time	Transition time	Transition quality
1	1	6	0	-
	2	12	0	-
	3	25	0	-
	4	37	0	-
[illegible]	5	6	[illegible]	[illegible]

Table 3.4 INVESTIGATED SCENARIOS [32]

The number of the test participants ought to be 25 and was consisting both from expert and non-experts views. Finally, all the viewers are prompted to change among the six provided channels as many times as they desire in order to form the best possible opinion so in the voting time of 10 seconds after each scenario to provide an accurate vote.

3.4.2 OBJECTIVE FRAMEWORKS

Alongside with the development of subjective frameworks, objective ones have been created. An example of objective framework it is being presented in [33]. Specifically, it is an evaluation method for video quality of full-reference type (FR). Furthermore, its outcome it is expressed with the assistance of Mean Opinion Score (MOS) scale [23], namely it is attempting to predict QoE as it would be calculated subjectively by a group of people with the use of MOS.

Moreover, this frameworks it is being described in four steps. In the first one, named "*Pre-Processing Stage*" is responsible for guarantee that so the source of video frames, as the received video frame are compatible for comparison. The next one is entitled "Calculating the perceptual difference of the aligned signals". This step reveals the observed indicators of video streaming. With the assistance of the third one, which is labeled "The classification of distortion parameters and indicators calculation", we are able to number the different factors that the received video streaming and video streaming sources have. Besides, a categorization can be realized based on which factors influence the most video quality. In the final step, entitled "Gaining the value of MOS", after the consideration and comparison of damaged parameters and various distortion factors, it is possible to estimate the value of the MOS.

Additionally, this method has the ability to check and confirm video quality the aid of Key Performance Indicators (KPI). Generally, video quality it is being affected by two elements. Those are network transmission and video content. In network transmission there are four indicators that influence the video quality. Those are:

- frame freeze/skips
- noise
- loss
- jitter

Besides, in video content there are three indicators:

- blockiness
- bluriness
- jerkiness

The main advantage of this method is the fact that it combines the two elements mentioned previously in order to acquire a comprehensive evaluation over the streaming video's quality. Hence the key indicators for this method are the following:

- *MOS:* This is the most significant indicator of these methods and evaluates video's quality in 1 - 5 scale, where 1 stands for the worst and 5 for the best.
- *Delay:* it is the elapsed time between the arrival of the last video streaming data frame and source video frame

- Distortion indicators:
- <u>*PSNR:*</u> it is categorized into three groups: PSNR Y, PSNR Cb, PSNR Cr. Moreover, it is utilized in order to analyze the quality in video streaming that are being received.
- <u>*Contrast:*</u> the outcome that will be arisen after the comparison between the "*receiving end video streaming data frame*" and the "*source of video frame*"
- <u>*Blur:*</u> it is the evaluation of video images about their clearness and their ability to accurately depict three-dimensional space distortion
- <u>*Frame skips and freezes:*</u> network overload outcome during inactivity, video packet loss or generally in any sort of harm.
- <u>*Brightness:*</u> "*the brightness value of the receiving end video streaming and video streaming source*"
- Blockiness: points the extent of the damage created by *"the lower stream rate and tough quantum of video images"*

Finally the result of this test indicates that the value of I-Frame Join Latency should not exceed the 2 seconds and that the MOS quality should be definitely over 3.6 (although it preferred to be greater than 4.0), otherwise would be unacceptable.

Another framework that uses objective metrics in order to evaluate QoE it is being presented in [34]. Specifically, this framework is able to evaluate QoE for video streaming which relies on the H.264/SVC codec. Moreover, it is the quantification of the main parameters of video stream, with the assistance of the SSIM and VQM full-reference metrics that determine the level of QoE. The four parameters that affect QoE in this framework are:

1. the video resolution
2. the scaling method
3. the video frame rate
4. the video content type

Additionally, for video clips it was utilized "*blue sky, crowd run and park joy*" in y4m format and having resolution of 1080p. Furthermore, with the assistance of JSVM software Version 9.15, videos where encoded in H.264/SVC having a variety of both spatial and temporal layers. The primal layer consists of 480*720 pixels, as resolution, whilst its frame rate is 1.875 fps (frames per second), while the maximum that video quality can get is a combination of 1216*684 pixels with 30 fps, which actually can be accomplished only if all the layers between the previously mentioned layers are available.

Finally, the results are being described in all cases are mapped to MOS values [23]. Firstly, are presenting the results concerning the different resolution levels. With the usage of the SSIM method it was observed that a decrease in resolution has a negative impact to the MOS values. Further, VQM method provides similar results, even though there are some differences between them such as the fact that this method's results are closer to the perceived quality that an end-user would experience. Finally, concerning the results for different frame rate, are being better predicted by VQM. Generally, the value of both VQM and SSIM metrics, it is being decreased while there is a lower frame rate. But in case where the value of the frame rate is 15 fps the corresponding value of SSIM in MOS scale would be 3, whilst of VQM would be 4, if it is not better.

CHAPTER 4

STANDARDIZATION ACTIVITIES

4.1 Introduction to standardization activities

Nowadays, Quality of Experience possesses a significant role in the transmission of any type of data (video, audio, etc) through internet, as the end-user's satisfaction is in the center of services providers' interest. Hence, a lot of attempts where realized from different organizations (e.g. ITU, ETSI, etc) by producing numerous standards or assessing the standardization activities in order to achieve to provide to the end-users the best possible experience.

One of the organizations that participate in standardization activities is the International Telecommunication Union (ITU). Specifically, in this organization there are different sectors, each one occupied with its subject. ITU-T (International Telecommunication Union, Telecommunication Standardization Sector) and ITU-R (International Telecommunication Union, Radiocommunication Sector) are in charge for the standardization of visual and audio methods. Moreover, under each sector there is a number of Study Groups (SG). In ITU-T sector the SG12 studies the performance, Quality of Service, and Quality of Experience of telecommunication services and is the main study group for these study items. Furthermore, SG12 is responsible for the standardization of various quality evaluation methodologies, especially for speech, which are being applied both in IP telephony services and in conventional PSTN/ISDN (public switched telephone network and integrated services digital network). Besides this study group investigates, for multimedia services (e.g. IPTV), QoE requirements and evaluation methods. Another study group which is under ITU-T is the SG9. This one is examines the provision of television services over cable and additionally studies quality evaluation methods for multimedia and video. The necessity for harmonization of the two study groups mentioned just above, lead to the creation of the Joint Rapporteurs' Group on Multimedia Quality Assessment (JRG-MMQA). Finally, SG6 which is under ITU-R is in charge for broadcasting services (e.g. radio services, television services). The QoE aspects of this study group are being studied in Working Party 6C (WP6C). [35]

Another organization that assists the development of standards is the Video Quality Experts Group (VQEG) which it was founded in 1997 by a group of experts on objective and subjective video quality derived from ITU-T and ITU-R. This organization acts like a technical advisory team, which is focused in objective quality assessment of video. Once a validation test, conducted by VQEG, is terminated, a final report it is being submitted to the ITU, which is the ultimate responsible for new standards concerning the evaluation of the objective perceptual quality. For example, there is cooperation between VQEG and SG9, as this study group utilizes the outcome of the VQEG's tests for its Recommendations.

There are also other organizations which are being involved in standardization activities. DSL forum has published a report for QoE requirements concerning the video transmission. Moreover, Video Services Forum (VSF) has recommended for video over IP transport-related metrics and also initiates an activity group on QoE metrics. Finally, at ATIS, the QoS Metrics Task Force within the IPTV Interoperability Forum is looking at QoE models for audio, video, multimedia and transactions. [36]

4.2 Standards for Video Quality

4.2.1 ITU-T P.910

In this Recommendation are being demonstrating subjective evaluation methods, which are non-interactive, and are appropriate for the one-way video quality measurement concerning multimedia applications (e.g. videoconferencing, tele-medical applications, etc). Moreover, these methods can be applied in numerous situations such as, for example, in the selection of algorithms.

Furthermore, in chapter 5 are being provided some instructions concerning the parts that compose the experiment. Specifically, information is given about the recording environment, the recording system (e.g. camera), and the characteristics that must have the scenes that would be shown during the test. The last one, has two crucial parameters; *“the spatial and temporal perceptual information”,* which are being analyzed.

Moreover, in this document are also being mentioned three test methods. The first one is the Absolute Category Rating (ACR) or the Single Stimulus Method. In this method the test sequences are being exhibited one at a time and after the scene under test is played, the test participants are asked to assess its quality. The test participants for the evaluation are using the following rating scale:

5 Excellent
4 Good
3 Fair
2 Poor
1 Bad

The next one is the Degradation Category Ration (DCR) or Double Stimulus Impairment Scale method. In this method, there is a pair of test sequences. The first one is the source reference and the second one, whilst the same, it is being exhibited via one of the systems which are under test. The test participants have up to ten seconds in order to evaluate the quality in second sequence, by judging its impairments, having in mind the quality that had the first one and using the following rating scale:

5 Imperceptible
4 Perceptible but not annoying
3 Slightly annoying
2 Annoying
1 Very annoying

The final is the Pair Comparison method (PC). In this one, as it can be easily understood by its name, the test sequences are presented in pair. Actually, it is the same sequence presented in two different systems, which are under test. First it is being presented in the one and later in the other, and afterwards the test participants have again up to ten seconds to evaluate.

Another thing that is being mentioned in this document concerning the conduction of an experiment is the number of participants. They must be at least four, whilst the maximum number is forty. Finally, there is a table in chapter 7, which includes the normal viewing conditions under which the video quality evaluation should be conducted. [31]

4.2.2 ITU-T J.bitvqm

The model described in this standard is an objective hybrid perceptual/bit-stream one and in contrary to the current video quality models which are provided processed video sequences, this model may have the possibility to access the transmitted bit stream video data. Because of this fact it can extract information about codec (e.g. type, bit-rates, frame rates, etc) and transmission errors (e.g. delay, packet loss, etc) and therefore it can be faster and more accurate. Currently ITU-T SG9 is responsible for the development of such models.

Moreover, it is known that if we measure the video quality every half second, it is possible to acquire sufficient information for processed video's sequence quality. Furthermore, with the assistance of video quality scores, an objective model like this one can improve significantly, especially for no-reference (NR) models. Finally, it is referred to this recommendation that the video quality scores can be transmitted as metadata or by using watermark techniques. [37]

4.2.3 ITU-T P.NAMS

In this standard it is being presented an objective parametric quality assessment model which has the ability to predict the impact that will have to the end-user's anticipation any potential IP network impairment, in any multimedia streaming and IPTV application. This model has the specificity to obtain information from packet headers only, namely its input parameter, and also manage to calculate the Mean Opinion Score (MOS) with prediction, on the ACR scale for audio and video parts of the stream. Moreover, this standard is not able to provide a comprehensive assessment over the perceived quality by the end-user as its scores reflect the impairments on the IP network which is being measured. In addition, as this model is a parametric one, and subsequently it has not access to the content of the IP packet, the scores that will predict would necessarily depict an average perceptual impairment and additionally may lead to deceptive results, as it may achieve high scores and yet the quality of the stream remains to a low level.

This standard has two distinct application areas. The first, which is the low bit-rate mode, is the QCIF-QVGA which is mostly for mobile TV and streaming. The second one, also known as the high bit-rate mode, is for SD and HD Television and mostly for IPTV. Furthermore, P.NAMS has four modes of operation as it can be implemented in different locations into the network and has as input different parameters. These four modes are:

1. Static operation
2. Non-embedded Dynamic operation
3. Distributed operation
4. Embedded operation

Finally, P.NAMS's output parameters are estimated ACR MOS, as well as separate video and audio ACR MOS. Moreover, optional parameters are jitter buffer size, bit rate (for both video and audio), mean packet size, delay variation range, etc. [38]

4.2.4 ITU-T J.143

This recommendation describes the end-user's requirements concerning the objective perceptual video quality measurements, so in digital cable television, as in similar applications. In digital television there have been produced new QoS considerations, which have complex relationships between objective parameters measurements and subjective picture quality. Specifically, objective measurements in combination with a good subjective quality assessment, is a main target so as to attain optimal Quality of Service in the operation of cable television systems.

Next in this recommendation are being referred the end-user's requirements. Firstly is the accuracy of perceptual video quality measurements. In this requirement it is being underlined the need of accurate measurement on the programme material that it is being delivered by cable television, independently of its content and of its impairments. Secondly is the availability of the input/reference signal. This case indicates the significance of having either the original input video or a reference signal, since the input video is unavailable, so as to be able to measure the perceptual video quality. Another requirement lies in case when a perceptual measuring device it is being positioned in the chain and should be operated on-line in a cable television system. Here, we do not want the insertion of the device in the chain to hinder the performance of the chain.

Also in this recommendation there is a reference to two distinct applications in which a cable television operator may want to perform an objective measurement of perceptual video quality. The first one is to test new equipment and the other is when wishes to install perceptual measuring devices along the programme chain.

Finally, it is referred that for the general good, all the manufacturers should share the details of their methods. Also, it is recommended that television operators should select the most suitable method (full reference, reduced reference, no reference) for their tests according to their needs.

The three methods mentioned just above, are being analyzed in the appendix I. [39]

4.2.5 ITU-T J.140

This recommendation mentions a subjective method for the evaluation of picture quality for digital cable television systems, from the signal source to the user’s receiver, including satellite links terrestrial links and/or cable links. Generally, the subjective evaluation methods are more suitable to assess the performance of a television system, because they use measurements that perceive more directly the reactions of the end-users.

Furthermore, in this recommendation it is being analyzed the test material that should be used during the tests. A number of approaches may be conducted in order to decide what kind of test material it is preferable in television assessments, nevertheless in practice particular kinds of test materials should be used to address a specific assessment problem. For example, the assessment problem of “overall performance with average material” it provokes when the material that is being used is “General, ‘critical but no unduly so' " (The example is taken from the table 1). Moreover, some parameters may lead to the same order of impairments for most pictures or sequences. In occasions like this one the results are collected with a very small number of pictures or sequences. Additionally, it is vital to include in the test material critical sequences, as it is possible to having this in mind when interpreting results.

Finally, in the Annex A it is being analyzed the SSCQE (Single Stimulus Continuous Quality Evaluation) method which has been developed by ITU-R and it is part of the ITU-R BT.500-7 Recommendation. [40]

4.2.6 ITU-T FG IPTV-C-0210

In this contribution, it is being addressed the QoE requirements and the relationship to QoS parameters. At first, there is a reference at video coding. Video information can be either encoded at the head end, with the assistance of video coding techniques (e.g. MPEG2, H.264) or at the content server. Furthermore, as IPTV supports both standard definition (SD) TV and high definition (HD) TV, there is an exhibition of the attributes of those two services. Additionally, in this contribution it is being mentioned the structure of an IPTV network which consists of four main network segments (the content server, the core network, the access network, and home network), as well as the priorities that should be held in the core network and the fact that the capacity of the access network plays key role on to how many channels are extended to the end user. Finally, it is being realized a short analysis to the transport impairments which are usually expressed in terms of packet delay and packet loss. Definitions are given for the packet transfer delay and packet loss ratio. [41]

4.2.7 ITU-T FG IPTV-IL-0006

In this Recommendation it is being described a method which is applicable only to those video services which have available a two-way digital communication. Specifically, it is a method more useful to the service provider, because offers to him the ability to reassemble the received video in order to examine the video quality at the receiver by using transmission error information for packetized video transmission.

In multimedia applications, at which the video data is being transmitted by using packets, a number of errors might take place just as packet loss and overflow, producing block errors, jitter, delay, etc in the received video. The digital communication has the advantage to exactly identify all this kinds of errors because the use of packets. So, if in digital video transmission, there are no errors from the provider to the end user, the quality of the video would be identical to both of them. Hence, if the receiver transmits a video which has error information (e.g. packet loss, delay) the provider is able to construct a duplicate of the video that it is being seen at the receiver. In the Table 1 it is being mentioned some of the most typical transmission error information. For example, if the type of transmission error is information on skipped or lost frames, then the contents of the transmission information would be skipped or lost frame indexes.

In order this method to achieve the things described below, transfer protocols such as Real-time Transport Protocol (RTP) and Asynchronous Transfer Mode (ATM) should be used and furthermore, this method requires an additional channel so that the receiver could be able to send transmission error information to the service provider. Finally, in the chapter 3 of the annex there is an analysis of all the types of messages that being transmitted for each error. [42]

4.2.8 ITU-T J.149

In this Recommendation it is being presented a number of methods in order to specify accuracy and cross-calibration of Video Quality Metrics (VQM). Specifically, in this recommendation there is an algorithm so that to quantify the accuracy a given VQM. This algorithm relies on statistical analysis relative to subjective data. Furthermore, there are methods for curve fitting VQM objective values to subjective data so as to ease the preciseness of the calculations and to generate a normalized objective value scale which will be used for cross-correlation among the different VQMs. Moreover, it is being referred a method, which for a given VQM plot the classification errors to determine the relative frequencies of “false differentiation”, “false tie”, “correct decision” and “false ranking”. In addition, this recommendation provides a simplified root mean square error calculation in order to quantify the accuracy of a VQM in case in which the subjective data has equal variance across the VQM scale.

Concerning the methods that are being mentioned in this document, it can be referred that their basis lies on subjective, as well as, objective assessment of the component video, e.g. like the one defined in the ITU-R Recommendation BT.601 and the methods that they use are similar to ITU-R BT. 500-11 Recommendation. Finally, if there is a reference to a data set for a VQM. It will be formed by objective values and mean subjective scores for various motion video sources (SRC) processed by a variety of hypothetical reference circuits (HRC). If there is a measurement which not belongs to the data set, the methods referred in this Recommendation have the ability to estimate with accuracy and cross-calibration for applications that are similar and within the scope of the defined data set. [43]

4.2.9 ITU-T J.341

In this Recommendation it is being presented an objective video quality assessment method, suitable for HDTV, when a full reference signal is available. Moreover, in the model description, referred in Annex A, it is mentioned that in order to predict this model uses psycho-visual and cognitive-inspired modeling, so as to simulate subjective perception. Additionally, in Annex A it is described in nine steps the score estimation procedure and exhaustively analyzed in clauses A1 through A9.

Some examples where the described method in this Recommendation can be applied, is at any potentially real-time in-service quality monitoring at the source, at lab testing of video systems, etc. Furthermore, it would be unfeasible to use this model to substitute the subjective testing, because correlation values between carefully designed and executed subjective tests, e.g. realized in two different laboratories, under normal circumstances their outcome will not be identical. Finally, another limitation for this model, concerns the frame freezing condition, as it was not validated for evaluating video in a re-buffering condition. [44]

4.3 Standards for Audio Quality

4.3.1 ITU-R BS.1534-1

In this standard it is being described a new method, called “MUSHRA” (MUlti Stimulus test with Hidden Reference and Anchor), which is suitable for the subjective assessment of intermediate audio quality. According to this document, subjective listening tests produce the most reliable results considering the measurement of quality in audio systems. In this new test method (MUSHRA), they aimed to give reliable and repeatable measure of systems having poor audio quality and based on the impairment scale of ITU-R BS.1116 (ITU-R BS.1534-1 inherits many aspects from this standard) would be positioned in its lower half. Furthermore, there are some instructions about the experimental design, such as the fact that the listening tests should be conducted in a way so the participants to the tests are not overloaded, something undesirable as it would lessen their accuracy of judgment. Moreover, the length of the sequences should be lower than 20 s for two reasons. First, this will lead to a more limited duration of the experiment and in this way the test participants are avoiding fatigue. Another fact that should be taken into consideration is that the test participants should be grade only one attribute at a time; otherwise they are getting confused and they do not give reliable results. In addition, although MUSHRA has a lot in common with ITU-R BS.1116 they differ in some other. For example in ITU-R BS.1116 the test participants can only compare each system with its reference signal, in contrary to the MUSHRA in which at any moment they can switch between the reference signal and any of the systems under test and consequently they can judge which system is more appropriate. Also, in MUSHRA the test participants tend to begin a session with rough estimation of quality, whilst in ITU-R BS.1116 they start with a detection process followed by a grading process.

In case in which it is observed an anomaly in the assigned scores, it is appropriate to note the events that produced the scores. Furthermore, the results of the experiment should be presented in layman terms, so that would be understandable from everybody. Details, also, should be presented so that a knowledgeable person can check empirically the outcome. Finally the results have to contain some standard information (e.g. description of the test material, number of assessors, etc). [45]

4.4 Standards for Speech Quality

4.4.1 ITU-T P.501

This recommendation presents a variety of test signals, with a range from low complexity to such signals with a high level of complexity which can be employed for numerous situations in telephonometry. These signals are divided in two categories, the “Non-speech-like signals” and the “Speech-like signals”. The first one of them is further divided into three following subcategories:

- Deterministic signals
- Random signals
- A combination of random and deterministic signals

The second category it is also further divided in the following categories:

- Composite Source signals
- Speech-like modulate noise
- Composed signals in frequency
- Complex composed signals

Additionally, in this document are being described technical signals, e.g. noise and sine waves. Besides of the reference that is being made to the test signals, in this recommendation are also referred the rules that should be obeyed whilst creating each type of test signal. Specifically, are being presented some significant properties such as probability density functions, density spectra, shaping filter responses, etc. After this, it follows a guideline concerning the normal application of the test signals without fully analyze each case. Finally, this Recommendation it is being accompanied by a CD-ROM, which includes the total of the test signals, so that there will not exist any case where the creation of a test signal could be done improperly. [46]

4.4.2 ITU-T P.564

In this document it is being described the minimum criteria that should be abided for objective speech quality evaluation models, which have the capability to anticipate the impact of observed IP network impairments on the one-way listening quality that an end-user experiences in IP/UDP/RTP-based telephony applications at 3.1 kHz for narrow-band telephony applications. In the Annex B, there is an extension for wire-band telephony at 7 kHz. Moreover, any models that are compliant with this recommendation predict Mean Opinion Scores (MOS) on the ACR listening scale. Additionally, the quality predictions performed by a model like this, are not relied on RTP stream that it is being analyzed, but assumes a normal voice payload and also it should take into consideration the voice codec. However, some impairments associated to the payload are not influence the final score (e.g. the acoustic background noise, the delay, the speech level, etc).Hence it is possible to acquire high scores, whilst the quality remains in low levels. Furthermore, it is not possible to supply an end-to-end assessment for the quality of transmission, as its results demonstrate only the impairments on the IP network which is under test.

The accuracy criteria referred in this Recommendation is the outcome of the comparison of a model's performance with the PESQ algorithm using the output mapping specific defined for the narrow-band telephony at 3.1 kHz. Moreover, these criteria are have the intention to avoid situations where would exist "false positive" or "false negative" errors. Finally, they criteria should be applicable to any device of a network, including the edge devices. [47]

4.4.3 ITU-T P.800

In this document are being presented methods and procedures which are performing subjective assessment in order to measure the quality of a transmission. Moreover, in this Recommendation are also being given instructions to the administrations for the conduction of subjective tests which they perform to their own facilities in order to measure the transmission quality. On contrary, here, there is no reference to those types of test that it is being already described in other ITU-T's Recommendations. Specifically at the following documents:

a) Determination of Reference and Relative Equivalents – see *Handbook on Telephonometry*, Geneva, 1993;

b) Determination of Loudness Ratings – see Recommendation P.78;

c) Determination of Articulation Ratings (A.E.N. values) – see *Handbook on Telephonometry*, Geneva, 1993.

Furthermore, the methods of this paper are believed to be appropriate to specify the levels of satisfaction that certain telephone connections are supposed to provide, as well as, it is supposed that they are able to support the different types of degradation factors (e.g. circuit nose, transmission errors, side tone, propagation time, etc).

At chapter 6 there are recommended methods. The first one is the "Conversation-opinion tests". In this method, it is attempted to reproduce in the laboratory, the service conditions that an end-user experiences in a telephone call. An extended description of this method exists in Annex A. Another one is the "Listening opinion tests", where the levels of realism are lower than the conversation tests, hence, the limitations are also less severe. Proposed tests methods for listening-only tests are the Absolute Category Rating (ACR), which is presented in Annex B. Additionally, another method that is being referred is the "Quantal-Response Detectability", which is analyzed in Annex C and describes tests that are appropriate for the assessment of threshold values and their connected probabilities. Furthermore, in Annex D it is being presented the "Degradation Category" method (DCR), which compares the testing system with an impeccable's quality standard reference signal and the perceived quality it is being evaluated in a five -point scale. A final method for listening-only tests is the "Comparison Category Rating" (CCR) that is being described in Annex E and it is a variation of DCR. In this method similarly with DCR, a comparison it is realized between the system under test and a reference signal, but the rating differs, as in this case the scale begin from "Much Better" and ends in "Much Worse", whilst in DCR begins from 5, where the degradation is noiseless and terminates in 1, where the degradation is irritating. Finally, another recommended method is the "Interview and survey", where transmission's quality can be specified by "service observations". This includes the questions that are asked to the interviewing test users and it is analyzed in Recommendation P.82. In order to be the outcome accurate, the minimum of the participants must be a hundred. Probably the main drawback of this method is the fact that a little control is provided over the detailed characteristics of the testing telephone connections. [23]

4.4.4 ITU-T G.107

In this recommendation it is being described a computational model, also known as E-model. It is a useful transmission planning tool as it can evaluate the transmission parameters that influence the conversational quality of handset telephony at 3.1 kHz. Therefore, the model being analyzed in this document could be used by transmission planners so as to reassure that the end-users will be satisfied by the end-to-end transmission performance. Moreover, it is mentioned that the most significant output from the model is the “Rating Factor” R, but it can be adjusted so that to provide estimates of end-user opinions. This kind of adjustments is made only to assist transmission planning purposes. Additionally, it is being referred that the E-model is not completely verified for the total of combinations of its input parameters. Thus, for combinations of high importance gives reliable results, but when it uses other parameters as input, the given output is unreliable.

As it mentioned above, the “Rating Factor” R is the most significant output of this model. Its composition is the following:

$$R = R_o - I_s - I_d - I_e\text{-}eff + A$$

Every of the parameters being in the right side of the equation it is being analyzed in the recommendation. Moreover, the parameters that composing the R have the following explanation:

Ro: Basic signal-to-noise ratio

Is: Simultaneous impairment factor

Id: Delay impairment factor

Ie: Equipment impairment factor

A : Advantage factor

Furthermore, in the first part of Annex A there is a number of examples for conditions in which the usage of E-model should be exercised with attention, e.g. “the overall level of the equipment impairment factors”, “the advantage factor A”, the “derivation methodology for new equipment impairment factors”, etc. In its second part, there is a number of situations in which the update of the earlier version allows the E-model to have better performance.

Finally, in Annex C there is the source code of the E-model, programmed in BASIC. [29]

4.4.5 ITU-T G.114

This Recommendation presents guidance on the influence of end-to-end one-way delay, also known as latency, and additionally it provides an upper bound in one-way network delay at 400 ms. Even though it is suggested that for general network planning the limit of 400 ms should not be overcome, in some cases of highly interactive tasks, such as video conferencing and interactive data applications, this limit judged as insufficient, because those tasks can be influenced by much lower delays. Moreover, if there are any delays below 500 ms, their effects on conversational speech can be calculated by a curve, which is derived from E-model, namely ITU-T’s G.107 Recommendation. The difference between the E-model and this Recommendation lies in the fact that this one supplies useful information concerning the one-way delay as a parameter by itself, whilst in G.107 (and its ITU-T Rec. G.108 and ITU-T Rec. G.109) the evaluation of the effects of the delay it is being made in combination with other impairments, e.g., distortions due to speech processing. Furthermore, for non-speech applications (e.g. interactive data, video, etc) there are not evaluation tools similar to E-model, so the influence that the delay would have on them must be carefully monitored. Additionally, as it referred previously, an upper bound of 400 ms should be kept. Nonetheless, this has and exceptions, e.g. an unavoidable double satellite hop for a hard-to-reach location. Finally, in Annex A it is being made a reference in the various occasions where a delay can be occurred, e.g. delay in wire-bound environment, delay in mobile and wireless environment, delay in IP environment (one frame per packet), etc. [48]

4.4.6 ITU-T P.851

In this document are being presented subjective evaluation methods which give information of the perceived quality by an end-user, concerning telephone services which are relied on spoken dialogue systems. It is also being described the assessment methods' various points of view concerning the perceived quality by an end-user and additionally it is being mentioned that the spoken dialogue system considered to be as a black box. The most significant of quality aspects are the following: "*the usability of the service, the communication efficiency, task and service efficiency, user satisfaction, perceived speech input and output quality, the system's cooperativity, the symmetry of the interaction, and the perceived smoothness of the interaction*".

Furthermore, the methods described in this document are relied on experiments conducted in laboratories, where the test participants have the possibility to interact with the examined spoken dialogue system so that to experience a realistic task. Moreover, the test participant's opinion concerning the perceived quality can be evaluated in three ways. The first one is with the use of questionnaires after the completion of the conducted experiment, which it is being analyzed in chapter 7. The other one is with a guided or an unguided way or, finally, with the assistance of other usability assessment methods. Finally, the set-up of such experiments it is being presented in chapter 6. [49]

4.4.7 ITU-T P.862

In this Recommendation it is being presented the objective method PESQ (Perceptual Evaluation of Speech Quality) which predicts the end-to-end speech quality of 3.1 kHz, namely narrow-band telephony and narrow-band codecs for speech. Moreover, in this document except from an excellent description of the method, it also provides instructions on how to use it and additionally it provides part of the outcome from a benchmark which was conducted by the Study Group 12 from 1999 since 2000. This method is essentially the extension of the PSQM, which was presented in the ITU-T P.861 Recommendation, which was only able to be used in the evaluation of speech codecs and on contrary could not support variable delay, short localize distortions and finally, filtering. PESQ is capable of supporting these effects with time alignment, time function equalization and an algorithm which calculates the average distortions over time and it is being described in chapter 10.

The way that PESQ functions is by the comparison of two signals. The first is the original signal X(t) and the second is a degraded signal Y(t). What this method predicts is the evaluation that would be made by test participants to Y(t), which represents the quality that is perceived by them, if a subjective listening test was conducted. Finally, in chapter 4, are being mentioned three tables which are essentially three categories of test factors, applications and coding technologies that influence PESQ. The first table includes those cases in which PESQ's outcome has acceptable accuracy (e.g. Transmission channel errors, Waveform codecs, Codec selection); the second one has those variables which when used in combination with PESQ the result provide us with unreliable predictions, whilst the third table has those factors, applications and coding technologies in which the method referred in this Recommendation has not been validated yet. [50]

4.4.8 ITU-T P.800.1

The main subject of this Recommendation is a new terminology which ought to be used in combination with speech quality expressions in terms of Mean Opinion Score (MOS). The reason for the creation of this terminology is to avoid the misunderstandings about whether specific values of MOS are related to talking quality, conversational quality or listening quality and additionally if they originate from network planning models, subjective tests or from objective models. [24]

Specifically, the abbreviation MOS can be accompanied with the following identifiers, which each one of them refers to a distinct area of application:

- N refers to narrow-band (300 – 3400 Hz)

- W refers to wide-band (50 – 7000 Hz)
- LQ refers to listening quality
- CQ refers to conversational quality
- S refers to subjective
- O refers to objective
- E refers to estimated

Finally, in this recommendation there are three categories, one for each different value of MOS. Specifically, the first is for listening-only situations:

1. MOS-LQS
2. MOS-LQO
    a. MOS-LQO (electrical)
    b. MOS-LQO (acoustical)
3. MOS-LQE

The second is for talking situations:

1. MOS-TQS
2. MOS-TQO
3. MOS-TQE

And the third and last one is for conversational situations:

1. MOS-CQS
2. MOS-CQO
3. MOS-CQE

4.5 Standards for Multimedia Quality

4.5.1 ITU-T J.148

In this standard it is being defined a number of requirements that is useful, in order to develop a multimedia perceptual model, and more precisely an objective one for auditory-visual services. Moreover, this model it is primarily used for the measurement of the quality of limited bandwidth services, such as services delivered at or below 2Mb/s. As it mentioned before, this model it is an objective one, so for reassure its validity, its results must be compared with quality ratings obtained from subjective tests. Also, this model consists of three input modules:

1. prediction of audio quality
2. prediction of video/composite image quality
3. indication of differential delay between audio and video

In addition, a fourth input permits the model to accept any task-dependent influences which may impact at the perceived quality. Furthermore, this model should be applicable both to the measurement of services that have a reference available and services that they have no reference information present. [51]

4.5.2 ITU-T G.OMVAS

This recommendation presents a computational model, suitable for video and audio-streaming, as well as for linear applications over IP, such as IPTV. This model is an assistant QoE/QoS planning tool, with which can be evaluated those video and audio parameters that affect the perceived QoE by the end-user. Its difference with the G.1070 recommendation is that, whilst the G.1070 is a full-duplex two-way video-telephony, this model supports one way video and audio streaming applications (e.g. IPTV). Moreover, its difference with the J.148 Recommendation lies in the fact that in this recommendation the quality it is being estimated by using network, application and terminal equipment parameters. Furthermore, this model can be a useful tool for the QoE/QoS planners, as it can reassure them that the end-users are content with end-to-end service quality.

In the Annex A of this model it is being mentioned a block diagram of G.OMVAS model and two tables, the first with the input parameters for quality-estimation module and the seconds with the input parameters for coefficient database. Additionally, there is a reference concerning the model's outputs, which are multimedia ACR MOS, with separate audio and video ACR MOS. Also, G.OMVAS must use different ACR MOS scale for different video format. Finally, there is a table which includes the assumptions of G.OMVAS, as well as a reference to G.OMVAS requirements in quality estimation accuracy, which is the same for all codecs and video formats. [52]

4.5.3 ITU-T FG IPTV-C-0411

In this contribution it is being proposed a set of metrics for QoE/QoS in IPTV, organized into three groups.

The first of them has perceptual *quality metrics,* which has the ability to provide high level scores for the quality that an end-user experiences (QoE) so in video, as in audio. Additionally, it offers immediate visibility of the impact that the vast number of impairments has to the end-user. Examples of this metric are the MOS-V, the estimated PSNR (EPSNR), the Video Service Transmission Quality (VSTQ), etc.

The second has the *video stream metrics*, which assist to the better comprehension of the performance and the configuration of the encoded video stream. It is being divided into two categories. In the first, the “video stream description” can be found information about the type of the codec being used, Group of Pictures structure and other key factors, such as length and image size. Examples of this metric are the codec type, GoP type, frames per second, etc. In the second category, the “video stream metrics”, it is given information about the proportion of different type of video frame that are impacted by packet loss and discard and to the overall video bandwidth. Examples of this metric are the Proportion of I frames impaired, the Proportion of P frames impaired, the Mean bandwidth, the Peak bandwidth, etc.

The final group is the one with the *transport metrics* and is divided into four categories. The first of them is the “Packet Loss Metrics”, gives significant information on IP packet loss before and after the effects of error correction. Examples of this metric are the Uncorrected Packet Loss Rate, the Packet Discard Rate, etc. The second one is the “FEC metrics”, which present the rate of the effectiveness of the FEC that it is being used. An example of this metrics is the FEC effectiveness. The next category is the “Reliable UDP metrics”, which inform about the performance of retransmission based protocols such as Reliable UDP. An example of this metrics is the proportion of packets retransmitted. Finally, it is the "Jitter and Delay metrics". Examples of this metric are the smoothing jitter, the MAPDV, etc. [53]

4.5.4 ITU-T G.1080

In this Recommendation are being specified the end-user requirements for QoE, concerning the IPTV. Those requirements are being defined from end-user's point of view and subsequently are unknown both in transport protocols and also in to the network deployment architectures. Furthermore, the specification of those requirements is for end-to-end services and also information is provided on how they affect the network transport and application layer behavior. Moreover, any compression coding schemes in this recommendation are examples, as well as, any detailed numeric values which is given as performance targets, e.g. bit rate, packet loss rate, etc. [54]

Specifically, this recommendation provides requirements for:

1. media compression and synchronization (QoE for video and audio)
2. channel zapping time (QoE for control functions)
3. VoD trick mode (QoE for control functions)
4. Metadata (QoE for other IPTV services)
5. browser (QoE for other IPTV services)
6. content navigation (QoE for other IPTV services)
7. Accessibility (this is for further study)

4.5.5 ITU-T FG IPTV-C-0354

In this contribution it is being made a proposal of three items to be considered in the "Quality of Experience requirements for IPTV". The first of those items is the EPG. The QoE requirements for this one are:

1. User-friendliness
2. Response time to display EPG page
3. Processing time for importing metadata

The second one is the metadata. Its QoE requirements are:

1. Availability
2. Data size

Finally, the third is the browser and its requirements are:

1. Characteristics of a television set
2. TV-like display
3. Document size
4. Character size
5. Navigation
6. Cookie
7. Other functions (they may optionally be considered as part of QoE criteria) [55]

4.5.6 ITU-T FG IPTV-C-0507

This contribution suggests two modifications to other documents. Specifically, to the FG-IPTV-DOC-0063 suggests adding another subsection 14 to the main index in order to define the QoE requirements for IPTV related to “Service Support”. Those requirements are the availability, the accessibility, the user friendliness and Response-time/Problem resolving time.

The second suggestion for modification is to add factor 3 to subsection 12.2, entitled “QoE requirements for metadata”, of section 12. Specifically the modification concerns the most critical aspect of metadata, the “Correctness”. [56]

4.6 Synopsis

A synopsis of the standards analyzed above can be depicted on the following tables. In the first one there is a presentation of the standards based on whether they are for video, speech, multimedia or audio. In the second one, the standards are grouped based on whether they describe an objective or subjective assessment methodology. Finally, there is a third one, which is the intersection of the first two tables.

Video	Audio	Speech	Multimedia
ITU-T P.910	ITU-R BS.1534-1	ITU-T P.501	ITU-T J.148
ITU-T J.bitvqm	-----------	ITU-T P.564	ITU-T G.OMVAS
ITU-T P.NAMS	-----------	ITU-T P.800	ITU-T FG IPTV-C-0411
ITU-T J.143	-----------	ITU-T G.107	ITU-T G.1080
ITU-T J.140	-----------	ITU-T G.114	ITU-T FG IPTV-C-0354
ITU-T FG IPTV-C-0210	-----------	ITU-T P.851	ITU-T FG IPTV-C-0507
ITU-T FG IPTV-IL-0006	-----------	ITU-T P.862	-----------
ITU-T J.149	-----------	ITU-T P.800.1	-----------
ITU-T J.341	-----------	-----------	-----------

Table 4.1 Video\Audio\Speech\Multimedia categorization

Objective	**Subjective**
ITU-T P.564	ITU-T P.800
ITU-T J.148	ITU-T P.910
ITU-T G.107	ITU-R BS.1534-1
ITU-T J.bitvqm	ITU-T J.140
ITU-T P.NAMS	ITU-T FG IPTV-C-0411
ITU-T J.143	ITU-T G.1080
ITU-T G.OMVAS	ITU-T P.851
ITU-T FG IPTV-IL-0006	-----------
ITU-T J.149	-----------
ITU-T J.341	-----------
ITU-T P.862	-----------

Table 4.2 Objective\Subjective categorization

	Video	Audio	Speech	Multimedia
Objective	• ITU-T J.bitvqm • ITU-T P.Nams • ITU-T J.143 • ITU-T J.341 • ITU-T FG IPTV-IL-0006 • ITU-T J.149		• ITU-T P.564 • ITU-T G.107 • ITU-T P.862 • ITU-T G.114	• ITU-T J.148 • ITU-T G.OMVAS
Subjective	• ITU-T P.910 • ITU-T J.140	• ITU-R 1534-1	• ITU-T P.800 • ITU-T P.851	• ITU-T G.1080

Table 4.3 The intersection of the first two tables

CHAPTER 5

Quality of Experience and applications

5.1 Application Areas for Quality of Experience

It is well known that the Quality of Experience relies on the context of use, which is to some extent, specified by the application domain. Therefore, it is suggested to take into account the targeted application domain whilst defining the Quality of Experience. Application domains are usually multifaceted ranging from unidirectional to bidirectional and multidirectional services having dissimilar content modalities like:

- Delivery (i.e. broadcast, streaming, and file) of various types of content, such as video, audio, etc
- Conversational or collaborative applications considering both arts and social aspects
- Medical, as well as, educational applications

According to the context of the application it could be feasible to utilize real-time calculation of Quality of Experience so as to take informed decisions about how to best use media and infrastructure resources. By having models for Quality of Experience it is easier the application planning process, as the performance it is being modeled in terms of user experience instead of just utilizing low-level metrics of performance, such as loss rates, delay, CPU, battery, as well as, memory usage. [1]

Following, there are some examples, which are mentioned in [1], describing some application areas:

1) ***Web and Cloud***: Nowadays, in internet-based applications the majority of traffic in computation and data it is being realized via the cloud. Nevertheless, latest approaches concerning the managing quality are mostly implemented within the domain of a single stakeholder. But the effectiveness of those approaches facing a problem, which is the inherent lack of information exchange among

the involved entities, considering network providers, infrastructure and end-users. In order to give a solution to this problem, flexible co-operation between the involved constituents finally enables each user to:

I. be able to use the offered cloud service
II. distribute content, interact and collaborate with other end-users within a dynamic, transparent and seamless way whilst optimizing Quality of Experience simultaneously

2) ***Multimedia Learning***: The science of multimedia learning depends on experimental comparison so as to calculate the level of achievement of the "*transfer learning*" results. The term "*transfer learning*" is mentioning to the capability of the learner to use the knowledge they gain so as to give a solution to a new problem. The most common metric for calculating the "*transfer learning*" is the "effect size". This indicates how many standard deviations of amelioration in transfer test performance were acquired as an outcome of the multimedia instructional feature under research.
3) ***Sensory Experience***: The use of multimedia applications is possible to stimulate also other senses like mechanoreception and olfaction. Hence, multimedia applications are enriched with sensory effects, such as scent, wind, ambient lighting effects and vibration, which are aligned with the actual multimedia applications and rendered on suitable devices, such as motion chairs, fans, ambient lights and scent vaporizers. An end-user will comprehend these extra sensory effects as it is offered to him or to her the impression that he or she is a part of a particular multimedia application. Thus, Quality of Experience has to go further of audio-visual considering every simple human senses, emotion and feelings of an end-user. This indicates that Quality of Experience is multi-sensorial, as well as, multi-dimensional.
4) ***Haptic Communication***: The current field of research on haptic communications has as target to enlarge conventional audio-visual communication unto presence in remote environments, manipulation and physical interaction in order to activate immersion.

Quality of Experience concerning haptic communication is still in a primary level, but nevertheless it is easy to comprehend that Quality of Experience concerning hearing and visual feedback has to be considered along with haptic feedback into a joint multi-sensorial and multi-dimensional 'Quality of Experience metric.

So, considering all the application areas which have been described previously, we can conclude that each application domain will probably demand different requirements in terms of Quality of Experience. Hence, it is important to supply specializations of a generally accepted definition of Quality of Experience concerning to the respective application domain taking into consideration its requirements which are formed by means of features and influence factors of Quality of Experience. Subsequently, in order to supply an application-specific definition for Quality of Experience it is necessary to collect those factors and features of Quality of Experience which are indicate the requirements of the application domain and include them to a generally accepted definition for Quality of Experience [1].

5.2 Cloud and Quality of Experience

As the core networks were evolved to be more fast and credible and as the broadband internet access were start to be used wide-spread it was observed a tendency to transfer numerous services from the end devices to remote data centers. This actually is what defines the term "Cloud Computing". Firstly, only those services which did not burden the delivery network, such as e-mail, were realized into the cloud, but nowadays a variety of services and applications are available so as the end users have access to them remotely. Ergo, this has led to increased demands on network's Quality of Service as the end-users are waiting higher standards to be met. [4]

Lately, a new kind of service it is being offered at cloud, which potentially have the most rigorous requirements on Quality of Service, the cloud gaming, which pair the notions of cloud computing and on-line gaming. Cloud gaming offers the whole game experience to the end-users remotely from a data center and consequently the players are not anymore dependent to a certain kind or quality of gaming hardware. On contrary, they have the possibility to use simple devices as long as they are connected to the internet and are able to display high definition (HD) video, as the service basically moves the processing power demanded to render a game away from end-user to a data center and streams back the whole game experience to the end-user as a high definition video. Usually, just the multiplayer games use the network, in which many end-users are connected to a server, which handles the game environment and accepts input commands and returns status updates. In this case, the number of exchanged data is normally a small one. On the other hand, in case of Cloud Gaming the total of an end-user experience has to be transferred via the network and this is where it is the main difference between the last and the typical On-line Gaming in terms of network Quality of Experience. Whilst in typical On-line Gaming the experience that an end-user perceives is produced at the client computer and hence network's performance does not affect the presentation, in Cloud Gaming it can severely influence the quality. Moreover, from a network point of view a lot of obstacles that has to be overtaken in order to manage a service like this one to correspond to in the quality that the end-users demand, as unlike to typical web services or to video streaming Cloud Gaming claims both low latency and high downlink bandwidth.

In [4] it is being realized an attempt so as to investigate those parameters which are relied on actual end-users perceptions in order to find the Key Influence Factors (KFI) concerning Quality of Experience in Cloud Gaming and in order to accomplish that, subjective user surveys are necessary.

So in [4] is described a local testbet which conducted at the University of Wurzburg and emulates a cloud gaming service, namely it offers to a test user an experience of a game identical to a cloud application.

Furthermore, at an IP network the connection is possible to be affected by a variety of factors, such as jitter, delay, packet re-ordering, packet duplication or packet loss, etc. Nevertheless, in case of cloud gaming, as it has been formed so far, QoE is only influenced by the following two parameters: packet loss and packet delay. More specifically, delay has an impact at the time that an end-user's action is realized and the outcomes that will be perceived by the end-user, namely the time that intervenes between the pressure of a button at a game-pad until the desirable action depicted on the screen. The other parameter, packet loss, has to do with the data packets which are rejected by the program as the software have a real-time restriction as it cannot wait for them to be transferred for a certain amount of time or in a specific order. Additionally, in [5] it is being mentioned that loss must be under 1%. From an end-user perspective, a case of lost of a late packet will result in similar quality degradation independently of the reason that provokes it.

Scenario ID	***Delay***	***Packet Loss***	***Direction***
B	0 ms	0.0 %	both
D1	80 ms	0.0 %	both
D2	200 ms	0.0 %	both
D3	300 ms	0.0 %	both
L1	0 ms	0.3 %	both
L2	0 ms	1.0 %	both
M1	40 ms	1.5 %	both
M2	180 ms	0.3 %	both
A1	120 ms	1.0 %	client to server
A2	120 ms	1.0 %	server to client

Table 5.1 Test Scenarios [4]

At Table 5.1 are being presented the different scenarios involved in the testbet and a brief description of them is following:

1) ***Scenario B***: in this case all parameters are zero.
2) Scenarios D1 to D3: these are delay only scenarios and each one of them has a different price for delay.
3) ***Scenarios L1 and L2***: scenarios with symmetric packet loss of 0.3% and 1%.
4) ***Scenarios M1 and M2***: mixed scenarios with both delay and packet loss.
5) ***Scenario A1***: scenario with asymmetric settings and on contrary to previous scenarios, is not bi-directional, but examines only the connection from client to server.
6) ***Scenario A2***: a scenario which has the same delay and packet loss prices, but examines only the connection from server to client.

Furthermore, at the following picture it is being presented a synopsis of the testbet including results for all the scenarios mentioned before. Specifically, it depicts MOS score for each scenario from the conducted survey.

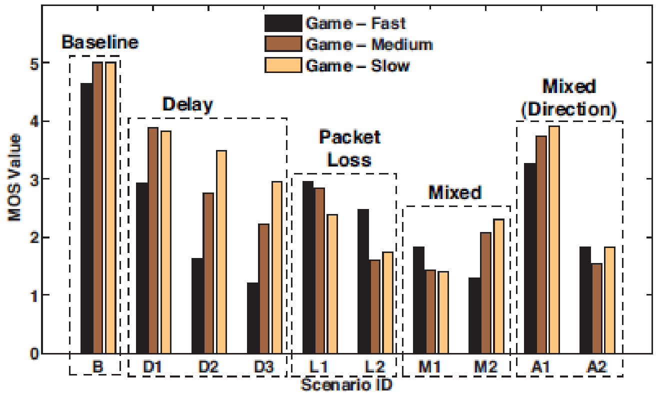

Figure 5.1 MOS Ratings per Scenario/Game [4]

The y-axis, in the Figure 5.1, indicates the MOS values whilst the x-axis includes all the scenarios. Moreover, for each scenario there are three bars which describe the type of game, namely whether it is fast or medium or slow.

Moreover, at Figure 5.2 are being presented the MOS scores for the bi-directional delay scenarios. Specifically, the scores are from B, D1, D2 and D3 scenarios and the conclusion that we come is that MOS values acquire smaller scores as the delay is increased and this means that the end-user anticipates poor QoE.

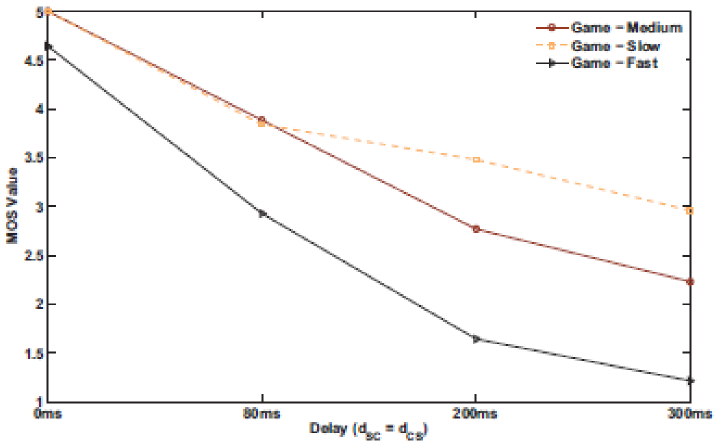

Figure 5.2 Delay Scenarios [4]

At Figure 5.3 is depicted how the end-user's perception of quality is influenced by packet loss. Again MOS values are indicated at y-axis, whilst those for packet loss at x-axis. The values for packet loss are taken from B, L1 and L2 scenarios. We can easily conclude that as the packet loss is augmented, the graphs at the picture headed towards to lower values and ergo the perceived quality by an end-user is poor.

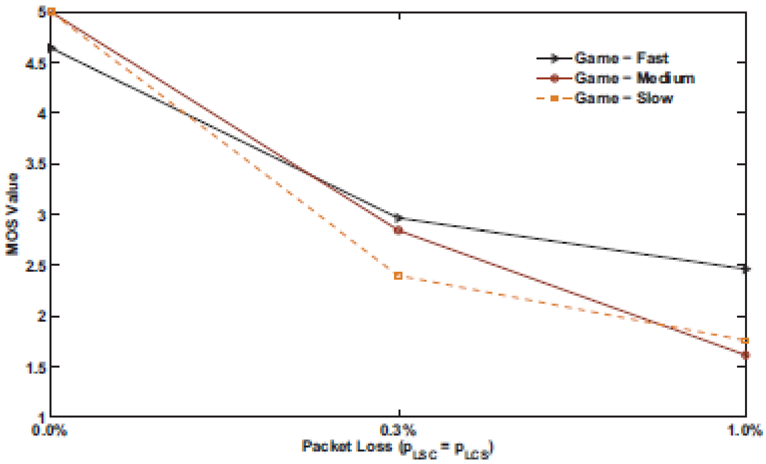

Figure 5.3 Loss Scenarios [4]

Next, from the Figure 5.4 and Figure 5.5 we can realize what impact more the end-user's perception concerning the quality of the service, the delay or the packet loss. At both figures the x-axis has the MOS values of M1 scenario while the y-axis has the M2 scenario MOS values. The color of the squares interpreted as follows. The darker the square is, the more the test users marked this combination of MOS score.

Specifically, at Figure 5.4, which is for medium game level, we can understand that there is a trend towards the M2 scenario which has less packet loss and more delay. So, we can conclude that in this type of game what mostly affects the experience that it is being perceived by an end-user is the packet loss.

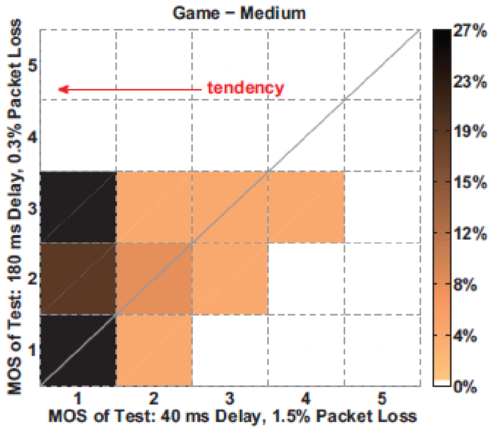

Figure 5.4 Mixed scenarios, Tendency to M2 scenario [4]

On contrary, at Figure 5.5, which describes a fast game level, there is a trend towards to M1 scenario, which has smaller delay than the previous one but it has higher packet loss. So, in this case the delay plays crucial role in the experience that will be perceived by an end-user.

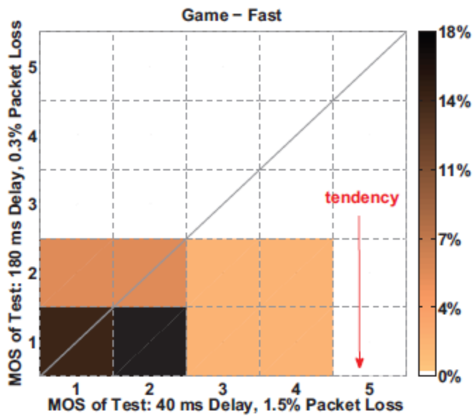

Figure 5.5 Mixed scenarios, Tendency to M1 scenario [4]

Finally, at Figure 5.6 is depicted the outcome of the asymmetric scenarios A1 and A2. To remind, in the first one, there is a disturbance to the connection from the client to server, whilst the reverse applies to the other. What derives from the graphs of the picture is that the disturbance of the server to client connection was impacted negatively to the perceived experience of the test users.

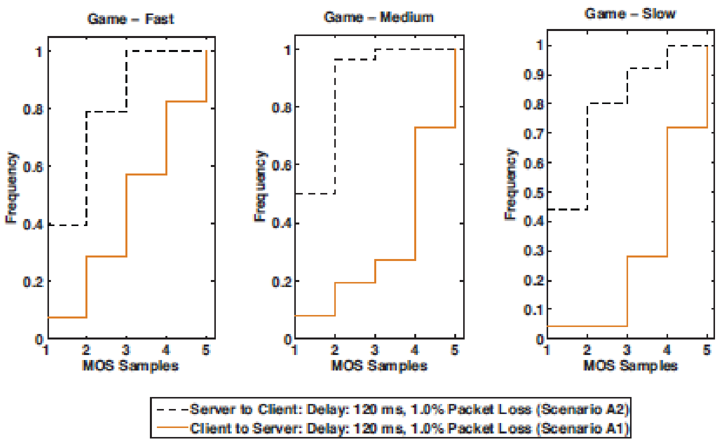

Figure 5.6 Asymmetric Scenarios [4]

Moreover, it is necessary for a service provider to be aware which parameter affects more the Quality of Experience regarding to the others, in order to construct and offer a service which could guarantee a constant minimum level concerning QoE. At the following Table 5.2 are being presented parameters and their weighted values. We can observe that the most significant parameters that influence the experience that perceives an end-user in cloud gaming is the downstream packet loss which corresponds to the weighted value 1. [4]

Parameter	***Weight***
Downstream Packet Loss	1.0
Downstream Delay	0.583
Upstream Packet Loss	0.370
Upstream Delay	0.212
Type of Game	0.067
Player Skill	0.006
Player Attitude Towards Game	0.006
Player Age	0.0

Table 5.2 Weight of parameters based on information gain [4]

Furthermore, in [59] it is being suggested an attempt to integrate cloud application into a QoE-based Content Distribution Network (CDN). CDN is evolved in order to ameliorate the quality of networks and to make even better the resource utilization. What actually happens in a CDN is the relocation of a server's content to other servers that are more near to the end-users, which are called "replica servers". As we can observe in the Figure 5.7, which depicts the suggested approach, the cloud provider shifts its data and services to a CDN provider and the last one is responsible for providing the data to the end-users. The focus in this approach is concentrated in two layers: the routing layer and the meta-routing layer. In the first one it is being suggested the QQAR (QoE Q-learning-based Adaptive Rooting) protocol, which is relied on a Q-learning algorithm mentioned in [60]. For the remaining layer, it is being suggested a server selection method which it is being presented in [61]. In both levels the feedbacks from end-users are essential for the QoE. [59]

5.3 IPTV and Quality of Experience

In its very first steps the internet was mostly utilized for services such as file transfer, e-mail and web-surfing. Nevertheless, during the past few years have been observed a constant growth in the use of internet for multimedia applications (i.e. VoIP services, such as Skype, IPTV, on-line gaming applications, etc) which on contrary to the traditional services (e.g. e-mail, web-browsing), are more delicate to bandwidth restrictions, jitter, packet loss and delay, and consequently any potential anomalies to these characteristics will cause great impact on the quality that an end-user will perceive and ergo they could influence the Quality of Experience. For example, a potential packet loss, even a small one, can cause visual artifacts for an IPTV service. [2]

IPTV is a collection of multimedia services which are being delivered via an IP network. Its main characteristics are:

1) Interactive television support
2) time shifting
3) low bandwidth demands
4) personalized content
5) accessibility with different kind of devices

As it is being depicted at the Figure 5.7, IPTV's network topology consists of five parts. These parts are analyzed underneath:

1) ***Head network***: it is from where the service provider makes available the video, to the IP network
2) ***Core network***: is responsible for the distribution of video flows from the header towards to the distribution networks of a service provider
3) ***Distribution network***: its bounds are from the backbone network until to the aggregation router, where the start of access network is.
4) ***Access network***: it allows the end-user to establish a connection with a service provider and permits the access of the multimedia content
5) ***Customer network***: make possible the communication and information exchanges from one computer to another and moreover enables the access to acquirable resources in the network.

Furthermore, IPTV it is being supplied by service providers which have to guarantee the Quality of Service by offering enough bandwidth and moreover have to guarantee the Quality of Experience .As it is being implied QoS is addressed more to an error-free transportation from provider's facilities to the end-user's computer through broadband wide-area network. On contrary, QoE has more to do with the general IPTV user experience, namely they will not suffer any potential service interruptions, zapping time that last long time or image degradations. In [3] there are referring four systems which have major responsibility for distributing IPTV services and are the following:

1) ***Video network header***: is the place where the content of programs and the applications are being kept.
2) ***Network***: is the mean which transports the interactive services and the TV content to the end-user.
3) ***Middleware***: is the software which manages the interactive services and the television content, all the way from the network header to the end-user's device via the network.

4) ***Customer's device***: namely the set-top-box, which is connected with the TV in the end-user's house.

Every one of these systems is able to influence the quality that an end-user will perceive and thus to affect the Quality of Experience. Despite the fact that there are a lot of subjective parameters that can affect QoE like *"content availability, easiness and available content indexation, user interface, colors palette, ergonomics, navigation design and program guide" [3]* there are two cases where QoE can be assessed objectively, which are the zapping measurements and the video/audio quality metrics. In the first one, the type of measures indicates the ability of the system to respond fast enough to the demand of the end-users to switch channels and acquire the right channel. A potential delay of one second and lower is reckoned as an acceptable time for zapping and hence the Quality of Experience for an end-user will be in a desirable level. In case of video/audio quality metrics, there is a majority of factors that can jeopardize the video and audio quality, such as the number of IPTV subscribers, the way they behave and the triple-play convergence. They are able to impact the perceived quality that an end-user will experience from an IPTV service. Moreover, any potential deterioration in the network like, latency, jitter, or loss packets and stream errors may also influence the Quality of Experience, as they can cause distortion, visual noise, etc. [3]

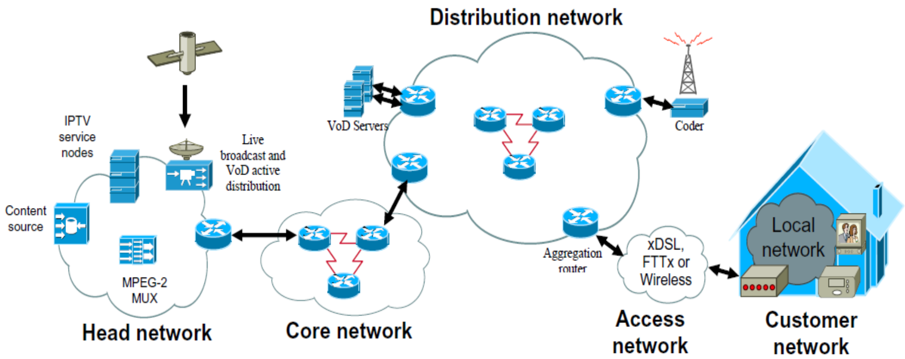

Figure 5.7 IPTV Infrastructure [3]

IPTV Quality Layer	***Metric***
Transaction Quality	IGMP join/leave latency, channel change delay, DRM error, Service availability error, service impairments error
Content Quality	MOS-V, MOS-A, MOS-AV, per PID bandwidth, program/channel bandwidth, PCR jitter, PCR failures count, ITF buffers overruns/underruns, AV synch
Media Stream Quality	TS sync loss count, sync byte error count, continuity count error count, Packet identifier error count, Presentation time stamp error count
Transmission Quality	RTP packet loss rate before/after error correction, RTP packet discard rate, Out of sequence packet rate, RTP burst loss rate before/after EC, Gap length, smoothing jitter, Peak packet to packet delay variation, RTP loss period count/loss distant count/minimum loss distance/maximum loss period, packet retransmissions

Table 5.3 ATIS IPTV Quality Metrics [9]

Moreover, in [62] there is a categorization of factors, from different international organizations, that influence QoE in IPTV service. At Figure 5.8 and at the Table 5.3 are presented those factors according to ATIS (Alliance for Telecommunications Industry Solutions) which suggests four IPTV quality layers and correspondingly quality metrics for each one of them.

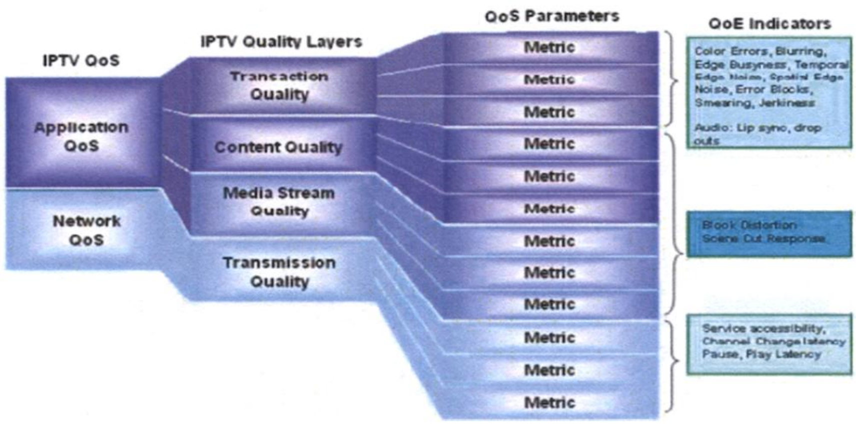

Figure 5.8 Classifying Quality Layers, QoS Parameters and QoE Indicators [62]

Next, DSL forum categorizes IPTV quality metrics as they appear at Table 5.4 and every layer can introduce an extensive view of Quality of Experience and also has metrics according to plane divisions. [62]

Layer	***Metric***	
Service Layer	Availability/Reliability/Survivability	
Application Layer	Control Plane	Channel change speed and scalability with load, VoD control, System start-up, EPG user interface navigation responsiveness
	Data Plane	Bit rate/ Resolution/ Application layer video encoding, Encoder quality and setting, Preprocessing, Tandem encoding and rate shaping
Transport Layer	Control Plane	IGMP processing time, Interleaving process time
	Data Plane	Loss, Delay, Jitter

Table 5.4 DSL Forum IPTV Quality Metrics [62]

Moreover, ITU-T's IPTV-GSI (Global Standard Initiative) has not have any different arrangement for IPTV QoS/QoE structure, instead they classify IPTV Quality of Experience layer as in Figure 5.9 that follows.

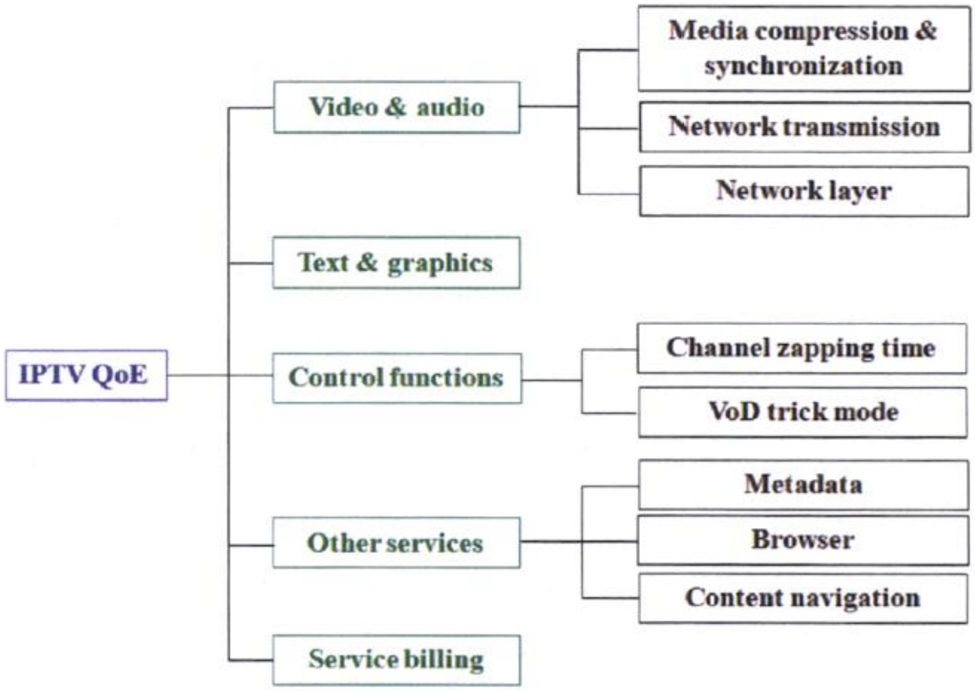

Figure 5.9 ITU-T's IPTV-GSI QoE Quality Layers [62]

CHAPTER 6

Conclusions and future challenges

The chapter presents a synopsis of all previous chapters and refers to future challenges of Quality of Experience.

6.1 Conclusion

The first chapter of this thesis firstly includes a reference to QoS that until recently was one of the few tools in the hands of researchers. In particular, QoS is defined as:

"QoS is a measure of performance at the packet level from the network perspective and performance of other devices involved in the service."

Furthermore, definitions for QoE are being mentioned stemming from different organizations such as ITU and ETSI:

QoE: “The overall acceptability of an application or service, as perceived subjectively by the end user.” (ITU-T)
Note 1: Includes the complete end-to-end system effects
Note 2: May be influenced by user expectations and context

Moreover, QoE is considered to be positioned above the Application layer of the OSI model since it is in higher abstraction level compared to QoS. Moreover, there is a brief analysis of the top three layers of OSI model, which are the most significant from a QoS/QoE point of view. Next, there is a reference to those aspects that affect QoE (e.g. interaction quality, usability, efficiency-related aspects, etc).

Additionally, there is a reference to those professionals (also called 'stakeholders'), that have interest of acquiring data concerning QoE, as they can improve their services and offer better products. Finally, at this chapter there is a summary of facts that we know about QoE so far.

The second chapter provides a definition of the two words that form the term "Quality of Experience", and afterwards there also is a variety of definitions for QoE. Additionally, a categorization of factors that influence QoE is being mentioned. These categories are the following:

1) Human Influence Factor (IF)
2) System Influence Factor (IF)
3) Context Influence Factor (IF)

Finally, this chapter describes the relation between QoS and QoE, along with the QoS parameters that are associated with the end-user.

The third chapter presents the evaluation methods for QoE, as well as the analysis of some frameworks that utilize these methods. Specifically, the first method is the subjective testing in which the participation of human testers is essential and it is probably the most reliable method for the evaluation of QoE. On the negative side, these methods are time consuming and costly. Moreover, objective methods come as an alternative to the time consuming and costly subjective testing, but according to some opinions they could be complementary to the last one in the evaluation of QoE. Objective methods are categorized in the following three groups based on whether they are use the original signal:

1) Full Reference (FR)
2) Reduced Reference (RR)
3) No Reference (NR)

Also, objective methods can alternatively, categorized and analyzed in this chapter and as following:

1) Media-layer models
2) Parametric packet-layer models
3) Parametric planning models
4) Bitstream layer models
5) Hybrid models

At chapter 4 the standardization activities concerning QoE are being presented, which were mostly conducted by ITU. Standards are being categorized based on the kind of service for which QoE is assessed. The kind of service could be:

1) multimedia
2) video
3) audio
4) speech

Finally, at the beginning of chapter 5 some application areas of QoE are mentioned and analyzed which are:

1) web and cloud
2) multimedia learning
3) sensory experience
4) haptic communication

Furthermore, at the second subchapter there is a reference to a specific game application in the cloud and how QoE is affected in this application. Lastly, at the third subchapter the QoE in IPTV service is presented.

6.2 Future challenges

As QoE is a relatively new notion is still evolving and ergo there are challenges that aim to optimize it.

In [26] is referred that late progresses on QoS and QoE has permitted the formation of a "*new ubiquitous wireless multimedia approach in the internet*". Wireless Mesh Networks (WMN) routing solutions with the assistance of QoE will be necessary in order to have successful multimedia communications, where the level of contentment for end-users from a specific service will be better than usual and mobile operators will have the possibility to augment their billing as operational cost will be minimized. Nevertheless, one of the most significant challenges is to evolve and implement novel in-service routing QoE solutions in WMNs.

Moreover, in [2] in which it is being described a QoE framework called "Perimeter", there is a reference to technical challenges which are:

- Even though simple weighted sum models appear to function quite well for calculating QoE relied on specific QoS parameters, there are other approaches which have fully achieve that. As the system will unstoppably create new inputs as it is being used, the assessment of QoE must be able to adjust to the new user inputs and thus to ameliorate over time.
- Definition of the P2P infrastructure so as to be possible to exchange QoE values between end-users.

Furthermore, a suggestion for future work at [65] it mentions that nevertheless the promising progress in the research which managed by the "convergence of the network and multimedia" by ameliorating the QoS/QoE, certain challenges are yet to be addressed, like "*the optimal resource allocation for large-scale multi-user systems and adaptable multimedia processing for triple-play mobile IPTV services*".

Additionally, at [66] are presenting QoE management challenges for cloud and multimedia cloud application as they appear to the Table 6.1 which follows.

Currently there are projects in progress which aim to further investigate and evolve QoE.

An attempt towards to this direction it is being realized by Queen (Quality of Experience Estimators in Networks) project which was initiated at September 2011 and will be by July 2014 [63]. The target of this project is to create multiservice QoE estimator agents which will stand for human users. These estimators relied on models of human perception and since a service is spotted, they will be trained progressively from the suitable profile for that specific service.

Finally, another project concerning QoE it is being conducted by IP Network Monitoring for Quality of Service Intelligence Support (IPNQSIS) project, which started October 2010 and will reach to its end at April 2013.

IPNQSIS's aim is to deploy continuous monitoring systems to examine the behavior of QoE via the analysis of network and service performance and the influence that will have to the end-users [64].

Topic	Topic QoE management challenge
QoE models for Cloud applications	Not yet mature and currently under research with a new scope of QoE for nonmedia services.
QoE monitoring and control mechanisms	Depend on the underlying QoE model, for estimating what, where and how to monitor. In an analogous way, QoE control mechanisms have to adequately react to performance issues, in order to maintain the desired QoE levels.
Overlay adaptation as a further step of QoE management for Cloud applications	Depends on the users' location and current situation in the Cloud and in the network with respect to traffic, available resources, etc.
Signaling between network and application to exchange information for QoE management	Requires new interfaces and network entities as discussed in the ALTO group.
Federation between clouds (similar to inter domain challenges of ISPs)	May be an inhibitor for QoE management, but needs to be realized with open interfaces an common standards.
Development and negotiations of SLAs	May provide the business fundamentals for QoE management.

Table 6.1 ***QoE management challenges for cloud and multimedia cloud applications.***

References

[1] European Cooperation in Science and Technology, COST 4146/10, Brussels, 2 July 2010.

[2] Eileen Dillon, Gemma Power, Manuel Ortiz Ramos, María A. Callejo Rodríguez , José Rodríguez Argente , Markus Fiedler, Dario S. Tonesi, " PERIMETER: A Quality of Experience Framework", research activity funded by the EU ICT FP7 project, Project No.: 224024.

[3] Miguel Garcia, Alejandro Canovas, Miguel Edo, Jaime Lloret, "A QoE Management System for Ubiquitous IPTV Devices", Third International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, 2009.

[4] ETSI TR 102 643 V1.0.1, Human Factors (HF); Quality of Experience (QoE) requirements for real-time communication services, 12/2009.

[5] Marcio Nieblas Zapater, Graca Bressan, "A proposed approach for Quality of Experience assurance for IPTV", Proceedings of the First International Conference on the Digital Society (ICDS'07) of IEEE, 2007.

[6] Ramesh Jain, "Quality of Experience", Published by the IEEE Computer Society, 2004.

[7] Bjorn Hestnes, Peter Brooks, Svein Heiestad, "QoE (Quality of Experience)- measuring QoE for improving the usage of telecommunication services", Telenor R&I Research Report R 21/2009, 24/09/2008.

[8] Cabral, R., in IFIP International Federation for Information Processing, Volume 252, Integration and Innovation Orient to E-Society Volume 2, eds. Wang, W., (Boston: Springer), pp. 193-199, 2007.

[9] Sebastian Moller, Klaus-Peter Engelbrecht, Christine Kuhnel, Ina Wechsung, Benjamin Weiss, "A taxonomy of Quality of Service and Quality of Experience of multimodal human-machine interaction", IEEE, QoMEX, 2009.

[10] http://www.thetanetworks.com/resources/qoe.html.

[11] Qualinet, Cost IC 1003, "Qualinet White Paper on Definitions of Quality of Experience", Version 1.0, 22 April 2012.

[12] http://www.techopedia.com/definition/25802/quality-of-experience-qoe ,Posted by: Cory Janssen.

[13] Dr. Raimund Schatz, Dr. Peter Reichl, "Quality of Experience – More Than Just Another Buzzword?" EuroView2011 Workshop, http://www.euroview2011.com/fileadmin/content/euroview2011/slides/slides_distribution_schatz.pdf, August 1-2 2011.

[14] http://searchcrm.techtarget.com/definition/Quality-of-Experience, Posted by: Margaret Rouse, last update in April 2008.

[15] http://en.wikipedia.org/wiki/Quality_of_experience ,last modified on 12 September 2012.

[16] "Amendment 2: New definitions for inclusion in Recommendation ITU-T P.10/G.100", ITU-T P.10/G.100.

[17] Rafal Stankiewicz, Andrzej Jajszczyk, "Quality of Experience for Multimedia in Clouds", IEEE COMSOC MMTC E-Letter.

[18] Stas Khirman, Peter Henriksen, "Relationship between Quality-of-Service and Quality-of-Experience for Public Internet Service".

[19] L. Liu, W. Zhou, J. Song, "The research of quality of experience evaluation method in pervasive computing environment", In Pervasive Computing and Applications, 2006 1st International Symposium on, pages 178–182, 2006.

[20] Hyun-Jong Kim, Dong-Hyeon Lee, Jong-Min Lee, Kyoung-Hee Lee, Won Lyu, Seong-Gon Choi, "The QoE Evaluation Method through the QoS-QoE Correlation Model", Fourth International Conference on Networked Computing and Advanced Information Management, 2008.

[21] P. Brooks and B. Hestnes, "User Measures of Quality of Experience: Why Being Objective and Quantitative is Important", IEEE Network, vol. 24, no. 2, Mar.–Apr. 2010, pp. 8–13.

[22] Rafal Stankiewicz, Piotr Cholda, Andrzej Jajszczyk, "QoX: What is it Really?", IEEE Communications Magazine, April 2011.

[23] "Methods for subjective determination of transmission quality", ITU-T Recommendation P.800.

[24] "Mean Opinion Score (MOS) Terminology", ITU-T Recommendation P.800.1.

[25] DSL Forum Technical Report TR-126 : "Triple-Play Services, Quality of Experience (QoE) Requirements", December 2006.

[26] Elisangela S. Aguiar, Billy A. Pinheiro, João Fabrício S. Figueirêdo, Eduardo Cerqueira, Antônio Jorge. G. Abelém, Rafael Lopes Gomes, "Chapter 6: Trends and Challenges for Quality of Service and Quality of Experience for Wireless Mesh Networks", from book "Wireless Mesh Networks", p.127-148, 2011.

[27] "Final text for G.RQAM, *"Reference guide to QoE assessment methodologies"* ", STUDY GROUP 12, Geneva, 18-27 May 2010.

[28] Akira Takahasi, David Hands, Vincent Barriac, "Standardization Activities in the ITU for a QoE Assessment of IPTV", IEEE Communications Magazine, February 2008.

[29] "The E-model: a computational model for use in transmission planning", ITU-T Recommendation G.107.

[30] Nicolas Staelens, Stefaan Moens, Wendy Van den Broeck, Ilse Mariën, Brecht Vermeulen, Peter Lambert, Rik Van de Walle, Piet Demeester, "Assessing Quality of Experience of IPTV and Video on Demand Services in Real-Life Environments", IEEE TRANSACTIONS ON BROADCASTING, VOL. 56, NO. 4, DECEMBER 2010.

[31] "Subjective video quality assessment methods for multimedia applications", ITU-T P.910.

[32] H. Fuchs, U. Jennehag, H. Thoma, "Subjective evaluation of low resolution tune-in streams for IPTV fast channel change", In Proceedings of IEEE International Symposium Broadband Multimedia Systems and Broadcasting, 2009.

[33] Bo Wang, Xiangmin Wen, Sun Yong, Zheng Wei, "A New Approach Measuring Users' QoE in the IPTV", Pacific-Asia Conference on Circuits,Communications and System, 2009

[34] Thomas Zinner, Oliver Hohlfeld, Osama Abboud, Tobias Hoßfeld, "Impact of Frame Rate and Resolution on Objective QoE Metrics", in Quality of Multimedia Experience, 2009, QoMEx 2009, International Workshop on, Trondheim, June 2010.

[35] Akira Takahashi," Framework and Standardization of Quality of Experience (QoE) Design and Management for Audiovisual Communication Services", NTT Technical Review, Vol. 7, No. 4, Apr. 2009.

[36] Stefan Winkler, Praveen Mohandas, " The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics", IEEE TRANS. BROADCASTING VOL. 54, NO. 3, SEPTEMBER 2008.

[37] "Terminal Requirements for Hybrid Perceptual/Bit-Stream Models", FG IPTV-C-0819.

[38] "A non-intrusive parametric model for the assessment of performance of multimedia streaming (P.NAMS)", ITU-T Draft Recommendation.

[39] ITU-T, "Recommendation J.143: User requirements for objective perceptual video quality measurements in digital cable television," May, 2000.

[40] "Subjective picture quality assessment for digital cable television systems", ITU-T Recommendation J.140 (1998).

[41] "Quality of experience (QoE), an overview", FG IPTV-C-0210.

[42] "Liaison statement to ITU-T FG IPTV - Q**o**S and Performance work for IPTV", FG IPTV-IL-0006.

[43] "ITU-T, "Recommendation J.149: Method for specifying accuracy and cross-calibration of Video Quality Metrics (VQM)". March 2004.

[44] "Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference", ITU-T J.341.

[45] "Method for the subjective assessment of intermediate quality level of coding systems", RECOMMENDATION ITU-R BS.1534-1.

[46] "Test signals for use in telephonometry", ITU-T P.501.

[47] "Conformance testing for narrowband voice over IP transmission quality assessment models", ITU-T P.564.

[48] "One-way transmission time", ITU-T G.114.

[49] "Subjective quality evaluation of telephone services based on spoken dialogue systems", ITU-T P.851.

[50] "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs", ITU-T P.862.

[51] "Requirements for an objective perceptual multimedia quality model", ITU-T J.148.

[52] "Opinion Model for video and audio streaming applications (G.OMVAS)," ITU-T Draft Recommendation.

[53] "IPTV QoS/QoE Metrics", FG IPTV-C-0411FG IPTV-C-0411.

[54] "Quality of Experience requirements for IPTV services", ITU-T FG IPTV G.1080.

[55] "A proposal on QoE of EPG, Metadata and Browser", FG IPTV-C-0354.

[56] "Suggested addition to Quality of Experience requirements for IPTV", FG IPTV-C-0507.

[57] Bart De Vleeschauwer, Filip De Turck, Bart Dhoedt Piet Demeester,Maarten Wijnants, Wim Lamotte, "End-to-end QoE Optimization Through Overlay Network Deployment".

[58] Michael Jarschel, Daniel Schlosser, Sven Scheuring, Tobias Hoßfeld, "An Evaluation of QoE in Cloud Gaming Based on Subjective Tests", Fifth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, 2011.

[59] Abdelhamid Mellouk, Said Hoceini, Brice Augustin, Nadjib Ait Saadi, Hai Anh Tran, Sami Souihi, "Adaptive Control based on Quality of Experience for Multimedia communications", IEEE COMSOC MMTC E-Letter, Vol. 6, No. 1, January 2011.

[60] Watkings and Daylan, "Technical note: Q-learning", Machine Learning, vol.8, No.3, pp.279-292, 1992.

[61] T.Lu, D. Pal, M. Pal, "Contextual multi-armed bandits", in Proceedings of the 13th International Conference on Artificial Intelligence and Statistics, 2010.

[62] Jaehyung Bae, Hong-Shik Park, Young Min Chin, Han-Choon Park, "An Identification of Correlation between Network QoS and Video Quality Index for IPTV Service", ICACT, February 15-18, 2009.

[63] http://www.celticplus.eu/Projects/Celtic-projects/Call8/QUEEN/queen-default.asp

[64] http://www.celticplus.eu/Projects/Celtic-projects/Call7/IPNQSIS/ipnqsis-default.asp

[65] Jinsul Kim, Tai-Won Um, Won Ryu, and Byung Sun Lee, Electrical and Telecommunications Research Institute Minsoo Hahn, Information and Communications University, "IPTV SYSTEMS, STANDARDS AND ARCHITECTURES: PART II, Heterogeneous Networks and Terminal-Aware QoS/QoE-Guaranteed Mobile IPTV Service", IEEE Communications Magazine, May 2008.

[66] Tobias Hoßfeld, University of Würzburg Raimund Schatz, Telecommunications Research Center Vienna — FTW Martin Varela, VTT Technical Research Centre of Finland Christian Timmerer, Alpen-Adria-Universität Klagenfurt, "QOE MANAGEMENT IN EMERGING MULTIMEDIA SERVICES Challenges of QoE Management for Cloud Applications", ", IEEE Communications Magazine, April 2012.