



Quality of Experience for Audio-Visual Services

Mamadou Tourad Diallo, Hassnaa Moustafa, Hossam Afifi, Khalil Ur Rehman
Laghari

► To cite this version:

Mamadou Tourad Diallo, Hassnaa Moustafa, Hossam Afifi, Khalil Ur Rehman Laghari. Quality of Experience for Audio-Visual Services. UP-TO-US '12 Workshop: User-Centric Personalized TV ubiquitOus and secUre Services, Jul 2012, Berlin, Germany. Fraunhofer FOKUS, 2012, pp.299-305. hal-01077548

HAL Id: hal-01077548

<https://hal.science/hal-01077548v1>

Submitted on 25 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Quality of Experience for Audio-Visual Services

Mamadou Tourad DIALLO
Orange Labs, issy
les moulineaux, France
mamadoutourad.diallo@orange.com

Hassnaa MOUSTAFA
Orange Labs, issy les moulineaux,
France
hassnaa.moustafa@orange.com

Hossam AFIFI
Institut telecom
Telecom SudParis
hossam.Afifi@it-
sudparis.eu

Khalil Laghari, INRS-EMT,
University of Quebec,
Montreal, QC, Canada
Khalil.laghari@gmail.com

ABSTRACT

Multimedia and audio-visual services are observing a huge explosion during the last years. In this context, users' satisfaction is the aim of each service provider to reduce the churn, promote new services and improve the ARPU (Average Revenue per User). Consequently, Quality of experience (QoE) consideration becomes crucial for evaluating the users' perceived quality for the provided services and hence reflects the users' satisfaction. QoE is a multidimensional concept consisting of both objective and subjective aspects. For adequate QoE measure, the context information on the network, devices, the user and his localization have to be considered for successful estimation of the user's experience. Up till now, QoE is a young subject with few consideration of context information. This paper presents a new notion of user's experience through introducing context-awareness to the QoE notion. A detailed overview and comparison of the classical QoE measurement techniques are given. The literature contributions for QoE measurement are also presented and analyzed according to the new notion of user's experience given in this paper.

General Terms:

Measurement, Human Factors,

Keywords:

QoE, objective, subjective, hybrid, MOS, Human behavior, human physiological, cognitive factors

Introduction

With the network heterogeneity and increasing demand of multimedia and audio-visual services and applications, **Quality of Experience (QoE)** has become a crucial determinant of the success or failure of these applications and services. As there is burgeoning need to understand human hedonic and aesthetic quality requirements, QoE appears as a measure of the users'

satisfaction from a service through providing an assessment of human expectations, feelings, perceptions, cognition and acceptance with respect to a particular service or application [1]. This helps network operators and service providers to know how users perceive the quality of video, audio, and image which is a prime criterion for the quality of multimedia and audio-visual applications and services [2]. QoE is a multidimensional concept consisting of both objective (e.g., human physiological and cognitive factors) and subjective (e.g., human psychological factors) aspects. As described in [3], context is fundamental part of communication ecosystem and it influences human behavior. Sometimes as human behavior is subjective and random in nature, it varies over the time. So, an important challenge is to find means and methods to measure and analyze QoE data with accuracy. QoE is measured through three main means: i) objective means based on network parameters (e.g. packet loss rate and congestion notification from routers), ii) subjective means based on the quality assessment by the users giving the exact user perception of the service, and iii) hybrid means which consider both objective and subjective methodologies. Some research contributions also present methods to evaluate the QoE based on users' behavior, technical parameters, some statistical learning, and measuring the users' feedbacks aiming to maximize user satisfaction and optimize network resources uses. Furthermore, there is a growing demand for methods to measure and predict QoE for multimedia and audio-visual services through some tools as (conviva, skytide, mediamelon etc)[4,5,6]. These tools use metrics such as startup time, playback delay, video freeze, number of rebuffering caused by network congestion to evaluate the perceived quality by the users and measure the QoE. According to [7], the parameters that affect QoE can be generally classified into three types: The quality of video/audio content at the source, the delivery of content over the network and the human perception (including acceptance, ambiance, preferences, needs, etc).

However, QoE as defined and measured today is not sufficient to adapt content or delivery for improving the users' satisfaction. Consequently, we define in this paper a new notion for QoE introducing more contextual parameters on the user and his environment to accurately predict the QoE. To maximize the end-user satisfaction, it is important to do some adaptation at both the application layer (e.g. choice of the compression parameters, change bitrate, choice of layers which will be send to client) and the network layer (e.g. delivery means, unicast, multicast, choice of access network, source server choice, delivery from a cloud...). The remainder of this paper is organized as follows, Section 2 defines a context-aware QoE notion, Section 3 gives a review on the QoE measuring techniques with a comparison between them, Section 4 describes the related research contributions on QoE evaluation mean comparing them according to the user experience notion. Finally, Section 5 concludes the paper and highlights some perspectives for the work.

2. Context-Aware QoE

For improving the QoE and maximizing the user's satisfaction from the service, we extend the QoE notion through coupling it with different context concerning the user (preferences, content consumption style, level of interest, location,) and his environment including the used terminal (capacity, screen size, ..) and network (available bandwidth, delay, jitter, packet loss rate..).

Human moods, expectations, feelings and behavior could change with variation in his/her context [8]. The context-aware QoE notion presents an exact assessment of QoE with respect to contextual information of a user in a communication ecosystem. To measure user experience, context information needs to be gathered as a first step in a dynamic and real-time manner during services access. This information includes: i): Devices Context (Capacity, Screen size, Availability of Network Connectivities ...), ii) Network Context (jitter, packet loss, available bandwidth ...), iii) User Context (Who is the user, his preferences, his consumption style, gender, age...), and iv) User Localization. User localization considers both the physical location of the user (indoor or outdoor) and the user location within the network that we call the Geographical Position Within the Network (GPWN) as illustrated in Figure 1. For the physical localization, user's can be localized indoors (within their domestic sphere for examples) through the Wi-Fi or Bluetooth techniques [9]. RFID can be used also to localize users with respect to RFID readers that are

supposed to have known places. For the outdoor localization of users, the following methods can be employed: GPS (Global Positioning System), Radio Signal Strength Indicator (RSSI), and Cell-ID (based on the knowledge of base stations locations and coverage). On the other hand, the GPWN localization of users aims to indicate the users' proximity to base stations, DSLAM, access points, content source servers...etc. This latter type of localization allows to choose in an adaptive manner the most suitable content delivery mean (for example, multicast can be used if many users are proximate to the same DSLAM, the content source server can be chosen according to its proximity to users, users who are near the base stations can receive the high video/audio quality for example, ...).

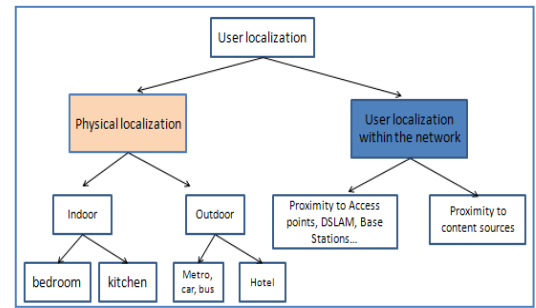


Figure 1: Localization principle

After context information gathering, a second step is to personalize and adapt the content and the content delivery means according to the gathered context information for improving user's satisfaction of services and for better resources consumption. Figure 2 illustrates our vision of content adaptation.

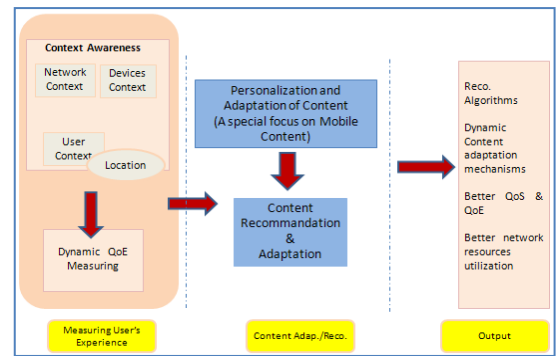


Figure 2: New vision of content adaptation

This paper focuses on the QoE measuring means presenting a comparison between them showing their suitability and/or limitation to the new notion of user's experience described in this section. The content personalization and adaptation are out of the scope of this paper and are a next step of our work.

3. QoE Measuring Techniques

Internet Service Providers (ISPs) use Quality of Service (QoS) parameters such as bandwidth, delay or jitter to guarantee good service quality. QoS is achieved if a good QoE is also achieved for the end users in addition to the classical networking configuration parameters [10]. The challenging question is how to quantify the QoE measure. In general there are three main techniques for measuring the QoE as discussed in the following sub-sections.

3.1 Objective QoE Measuring Techniques

Objective QoE measuring is based on network related parameters that need to be gathered to predict the users' satisfaction. Objective QoE measuring follows either an intrusive approach that requires a reference image/video/audio to predict the quality of the perceived content or a non intrusive approach that does not require reference information on the original content.

3.1.1 Intrusive Techniques

A lot of objectives QoE measurement solutions follow an intrusive approach. They need both the original and degraded signal (audio, video, and image) to measure QoE. Although intrusive methods are very accurate and give good results, they are not so feasible in real-time applications as it's not easy to have the original signal. The following subsections present some objective intrusive techniques.

- **PSNR (Peak Signal-to-Noise Ratio)** is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. It is defined via the Mean Squared Error (MSE) between an original frame o and the distorted frame d as follows [11].

$$MSE = \frac{1}{M \cdot N} \sum_{m=1}^M \sum_{n=1}^N |o(m, n) - d(m, n)|^2$$

Where each frame has $M \times N$ pixels, and $o(m, n)$ and $d(m, n)$ are the luminance pixels in position (m, n) in the frame. Then, PSNR is the logarithmic ratio between the maximum value of a signal and the background noise (MSE). If the maximum luminance value in the frame is L (when the pixels are represented using 8 bits per sample, $L = 255$) then

$$PSNR = 10 \log \frac{255^2}{MSE}$$

- **Perceptual Evaluation of Video Quality (PEVQ):** PEVQ is an accurate, reliable and fast video quality measure. It provides the Mean Opinion Score (MOS) estimates of the video quality degradation occurring through a network, e.g. in mobile and IP-based networks. PEVQ can be ideally applied to test video telephony, video conferencing, video streaming, and IPTV applications. The degraded video signal output from a network is analyzed by comparison to the undistorted reference video signal on a perceptual basis. The idea is to consider the difference between the luminance and the chrominance domains and calculates quality indicators from them. Furthermore the activity of the motion in the reference signals provide another indicator representing the temporal information. This indicator is important as it takes into account that in frames series with low activity the perception of details is much higher than in frames series with quick motions. After detecting the types of distortions, the distorted detected information is aggregated to form the Mean Opinion Score (MOS) [12].
- **Video Quality Metric (VQM):** VQM is a software tool developed by the Institute for Telecommunication Science (ITS) to objectively measure the perceived video quality. It measures the perceptual effects of video impairments including blurring, jerky/unnatural motion, global noise, block [13].
- **Structural Similarity Index (SSIM):** SSIM uses a structural distortion based measurement approach. Structure and similarity in this context refer to samples of the signals having strong dependencies between each other, especially when they are close in space. The rationale is that the human vision system is highly specialized in extracting structural information from the viewing field and it is not specialized in extracting the errors. The difference with respect to other techniques mentioned previously such as PEVQ or PSNR, is that these approaches estimate perceived errors on the other hand SSIM considers image degradation as perceived change in structural information. The resultant SSIM index is a decimal value between -1 and 1, where the value 1 indicates a good score and the value -1 indicates a bad score [14].

3.1.2 Non Intrusive Techniques

It is difficult to estimate the quality in the absence of the reference image or video which is not usually available all time, as in

streaming video and mobile TV applications. The objective non intrusive approach presents methods that can predict the quality of the viewed content based on the received frames without requiring the reference signal but using information that exist in the receiver side. The following are some methods that predict the user perception based on the received signals.

The method presented in [15] is based on the blur metric. This metric is based on the analysis of the spread of the edges in an image which is an estimated value to predict the QoE. The idea is to measure the blur along the vertical edges by applying edge detector (e.g. vertical Sobel filter which is an operator used in image processing for edge detection.). Another method is presented in [16] based on analyzing the received signal from the bit stream by calculating the number of intra blocks, number of inter blocks, and number of skipped blocks. The idea proposed in this work is to predict the video quality using these parameters. The predictor is built by setting up a model and adapts its coefficients using a number of training sequences. The parameters used are available at the decoder (client side).

The E-model proposed in [17] uses the packet loss and delay jitter to quantify the user perception of service. The E-model is a transmission rating factor R:

$$R = R_o - I_s - I_d - I_e + A,$$

Where R_o represents the basic signal-to-noise ratio, I_s represents the impairments occurring simultaneously with the voice signal, I_d represents the impairments caused by delay, and I_e represents the impairments caused by low bit rate codecs. The advantage factor A is used for compensation when there are other advantages of access to the user.

3.2 Subjective QoE Measuring Techniques

Subjective QoE measurement is the most fundamental methodology for evaluating QoE. The subjective measuring techniques are based on surveys, interviews and statistical sampling of users and customers to analyze their perceptions and needs with respect to the service and network quality. Several subjective assessment methods suitable for video application have been recommended by ITU-T and ITU-R. The subjective measures present the exact user perception of the viewed content (audio, video, image...) which is considered as a better indicator of video quality as it is given by humans.

The most famous metric used in subjective measurement is the MOS (Mean Opinion Score), where subjects are required to give a rating using the rating scheme indicated in Table 1.

In order to analyse subjective data, quantitative techniques (e.g., statistics, data mining etc) and qualitative techniques (e.g., grounding theory and CCA framework) could also be used [18]. Once subjective user study is complete, data are to be analyzed using some statistical or data mining approaches. Conventionally, non-parametric statistics is used for ordinal and nominal data, while parametric statistic or descriptive statistics is used for interval or ratio data.

Table 1: MOS Rating (source ITU-T)

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

3.3 Hybrid QoE Measuring Techniques

Hybrid QoE measurement merges both objective and subjective means. The objective measuring part consists of identifying the parameters which have an impact on the perceived quality for a sample video database. Then the subjective measurement takes place through asking a panel of humans to subjectively evaluate the QoE while varying the objective parameters values. After statistical processing of the answers each video sequence receives a QoE value (often, this is a Mean Opinion Score, or MOS) corresponding to certain values for the objective parameters. To automate the process, some of the objective parameters values associated with their equivalent MOS are used for training an RNN (Random Neural Network) and other values of these parameters and their associated MOS are used for the RNN validation. To validate the RNN, a comparison is done between the MOS values given by the trained RNN and their actual values. If these values are close enough (having low mean square error), the training is validated. Otherwise, the validation fails and a review of the chosen architecture and its configurations is needed [19]. PSQA (Pseudo-subjective Quality Assessment) is a hybrid technique for QoE measurement and is illustrated in Figure 3. In PSQA, training the RNN system is done by subjective scores in

real-time usage. The system maps the objective values to obtain the Mean Opinion Score (MOS). The advantages of this method are minimizing the drawbacks of both approaches as it is not time consuming and it does not require manpower (except in the subjective quality assessment preliminary step).

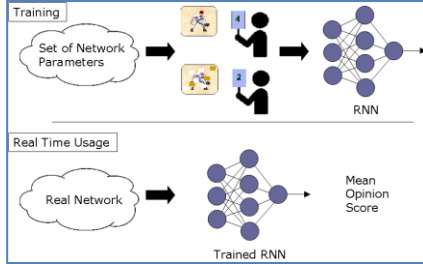


Figure 3: PSQA principle

3.4 Comparison of QoE Measuring Techniques

We compare the previously discussed QoE measurement techniques through considering the required resources, feasibility, accuracy and application type. Table 2 illustrates this comparison.

Table 2: Comparison of QoE measuring means

parameters		REQUIRED RESOURCES	FEASIBILITY	ACCURACY	APPLICATION TYPE
methods					
OBJECTIVES	INTRUSIVE	Gathering network parameters requiring more signaling, monitoring sensors, monitoring algorithms	It's easy to collect network parameters and to have reference image/video for no real time traffic and then measure the QoE	Does not consider users' opinions. Not very accurate	Not suitable for Real time traffic (image/VoD)
	NON INTRUSIVE	High receiver (user terminal capacity)	Less useful in real applications as in mobile equipment where the computational power frequently is limited.	Does not consider users' opinions. Not very accurate	Suitable for Live traffic (Streaming videos, Mobile TV)
SUBJECTIVE		Manpower Engaged users	Not realistic to ask all viewers to give scores	Very accurate (based on user's perception)	Any multimedia traffic
HYBRID		Engaged users. Random Neural Network (RNN)	Easy to have subjective measures (only once for learning). Once the tool has been trained, non linear function that can map any possible value of parameters into MOS.	Very accurate (subjective+objective)	Any multimedia traffic

4. Research Contributions on QoE

Measurement

Several methods to measure QoE exist in the literature which can be grouped into three main types: i) User behavior and technical parameters, ii) Learning process and iii) Measuring QoE based on subjective measure, objective and localization.

4.1 User behavior and technical parameters

The solution proposed in [20] predicts the QoE based on user engagement and technical parameters. This method quantifies the QoE of mobile video consumption in a real-life setting based on user behavior and technical parameters to indicate the network and video quality. The parameters are: the transport protocol, quality of video source, types of network that was used to transmit the video, number of handovers during transmission and the user engagement (percentage of the video that was actually watched by the user). Then, a decision tree is defined to give the Mean Opinion Score (MOS) based on these parameters.

4.2 Learning process: The learning process is based on collecting parameters and studying their behavior. It is based on linear regression, non linear regression, statistical methods, neural network tools etc... to derive weights for different parameters. Regression is the process of determining the relationships between the different variables. The method in [21] uses statistical learning to predict the QoE. After gathering the network parameters and considering linear regression, this technique calculates the QoE. The network parameters considered in this method are: packet loss rate, frame rate, round trip time, bandwidth, and jitter and the response variable is the perceived quality of experience (QoE) represented by (MOS). In [22], another method based on learning use combination of both application and physical layer parameters for all content types. The application layer parameters considered are the content type (CT), sender bitrate (SBR) and frame rate (FR) and the physical layer parameters are the block error (BLER) and a mean burst length (MBL). After gathering these parameters (in application and physical layer), non-linear regression is used to learn the relation between these parameters and MOS (Mean Opinion Score). An analytical formula is then used to predict the MOS.

Regression

4.3 Measuring QoE based on subjective measure, objective and localization

The method proposed in [23] considers the multi-dimensional concept of QoE by considering the objective, subjective and user localization to estimate the quality of viewed content. For gathering the subjective measure, after watching a video, the user answers a short questionnaire to evaluate various aspects of the watched video considering: the actual content of video, the picture quality (PQ), the sound quality (SQ), the emotional satisfaction. In addition, users are asked to which extend these aspects influenced his general QoE. During the video watching, objectives parameters (packet loss, jitter) are gathered by using RTCP Receiver Records (RR) these reports are sent each 5 seconds to the streaming server. To localize the user, the RSSI is recovered by the tool Myexperience. After calculating the correlations between parameters and between parameters and general QoE using the Pearson coefficient (most commonly used method of computing a correlation coefficient between variables that are linearly related), a general QoE measure can be modeled by using the multi-linear regression.*

4.4 Comparison of the Literature Contributions on QoE Measurement

We compare in Table 3 the previously discussed QoE measurement proposed in the literature through considering the network type, the localization, and the context consideration.

Table 3 Comparison of Literature Contributions

Techniques Parameters	Learning		User behavior and technical's parameters	Subjectives/ psychological and cognitives aspects	Objectives		Subjectives objectives and localization
	Regression based on applications and physicals parameters	Linear regression based on network parameters			Non intrusive	Intrusive	
Oriented	Mobile	Mobile	Mobile/fixed	Mobile/fixed	Mobile/fixed	Mobile/fixed	Mobile
User localization	NO	NO	NO	NO	NO	NO	YES
User preferences	NO	NO	NO	YES	NO	NO	YES
Dynamic consideration of network context	YES	YES	YES	YES	NO	YES	YES
Terminal context	NO	NO	NO	NO	NO	NO	NO
Considering the content context	NO	NO	NO	NO	NO	NO	NO

As we can see, there are lot of methods in the research field for measuring QoE. The learning methods include only network

context information without considering other context information as (device context, user localization, his preferences, content format...).

The method presented in [20], which is based on user behavior and technical parameters focuses only on the transport protocol, network parameters and user engagement. In this method there is no consideration of other context information like the user localization, and the terminal context. The presented technique in [23] uses much context information in QoE model. The user localization, the users' preferences, the network context are considered but the terminal context and the content context are omitted. Based on this comparison, our proposition is to build a QoE model which considers more context information to perform the QoE measure. Figure 4 illustrates our vision.

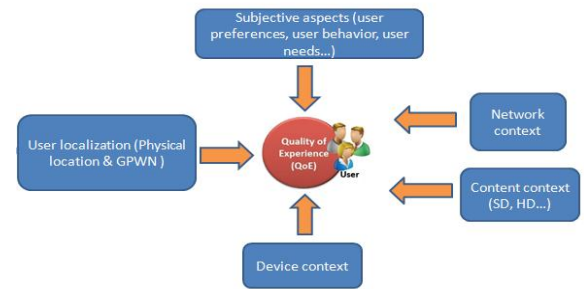


Figure 4: QoE Model

5. Conclusion and Perspectives

With the explosion of multimedia and audio-visual services and the high competition between service providers offers, Quality of Experience (QoE) becomes a crucial aspect for service providers and network operators to continue gaining users' satisfaction. In this context, there is burgeoning need to understand human hedonic and aesthetic quality requirements. This paper presents a study on the QoE measuring means considering both the classical methods and the research contributions. Classical methods consist of objective, subjective and hybrid techniques. A comparison is given between these methods based on the required resources, the feasibility, the accuracy and the application type suitability. The existing research contributions are also presented and classified into three types while presenting a comparison between them based on our presented vision of user experience. A new notion of QoE is presented in this paper considering context information (network context, device context, user context, content context...)

aiming to better improve the user satisfaction from the consumed service.

The perspective of this work is to associate the different context information with QoE measurement in a dynamic manner to satisfy the new notion of the user experience for content and delivery adaptation.

References

- [1] Laghari, Khalil Ur Rehman, Crespi, N.; Molina, B.; Palau, C.E.; , "QoE Aware Service Delivery in Distributed Environment," Advanced Information Networking and Applications (WAINA), 2011 IEEE Workshops of AINA Conference on , vol., no., pp.837-842, 22-25 March 2011.
- [2] T. Tominaga, T. Hayashi, J. Okamoto, A. Takahashi, "Performance comparison of subjective quality assessment methods for mobile video", June 2010
- [3] Laghari, Khalil ur rehman., Connelly, K., Crespi, N.: K. Laghari, "Towards Total Quality of Experience: A QoE Model in Communication Ecosystem". IEEE Communication Magazine April 2012.
- [4] www.conviva.com
- [5] www.skytide.com
- [6] www.mediamelon.com
- [7] F. Kuipers, R. Kooij, D. D. Vleeschauwer, K. Brunnström, "Techniques for Measuring Quality of Experience", 2010
- [8] V. George Mathew. (2001) ENVIRONMENTAL PSYCHOLOGY. <http://www.psychology4all.com/environmentalpsychology.htm>.
- [9] Deliverable D3 1 1 Context-Awareness- Version 15 Dec 2011.pdf, Task 3.1 (T3.1): Context Definition, Gathering and Monitoring
- [10] I. Martinez-Yelmo, I. Seoane, C. Guerrero, "Fair Quality of Experience (QoE) Measurements Related with Networking Technologies", 2010
- [11] A. N. Netravali and B. G. Haskell, "Digital Pictures: Representation, Compression, and Standards (2nd Ed)," 1995.
- [12] <http://www.pevq.org/video-quality-measurement.html>
- [13] M. Pinson, S. Wolf, "A New Standardized Method for Objectively Measuring", 2004 IEEE. Transactions on Broadcasting,
- [14] Z. Wang and Q. Li, "Video quality assessment using a statistical model of human", Journal of the Optical Society of America, 2007
- [15] P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi, "A no-reference perceptual blue metric", 2002
- [16] A. Rossholm, B. Lovstrom, "A New Low Complex Reference Free Video Quality Predictor", Multimedia Signal Processing, IEEE 10th Workshop, 2008
- [17] L. Ding and R. A. Goubran, "Speech Quality Prediction in VoIP Using the Extended E-Model", in IEEE GLOBECOM, 2003
- [18] Laghari, Khalil ur rehman., Khan, I., Crespi, N. K. Laghari, "Qualitative and Quantitative Assessment of Multimedia Services in Wireless environment" *MoVid, ACM Multimedia system USA* 22-24 Feb 2012.
- [19] K. D. Singh, A. Ksentini, B. Marienval, "Quality of Experience measurement tool for SVC video coding", ICC 2011
- [20] De Moor, K.; Juan, A.; Joseph, W.; De Marez, L.; Martens, L. "Quantifying QoE of Mobile Video Consumption in a Real-Life Setting Drawing on Objective and Subjective Parameters", , Broadband Multimedia Systems and Broadcasting (BMSB), 2011 IEEE International Symposium on IEEE, June 2011
- [21] Muslim Elkotob; Daniel Granlund; Karl Andersson; Christer Åhlund, "Multimedia QoE Optimized Management Using Prediction and Statistical Learning", on Local Computer Networks, IEEE, 2010
- [22] A. Khan, L. Sun, E. Ifeachor, J. O. Fajardo, F. Liberal Michal Ries, O. Nemethova, M. Rupp, "Video Quality Prediction Model for H.264 Video over UMTS Networks and their application on mobile video streaming", IEEE International Conference on Communications (2010)
- [23] István Ketykó, Katrien De Moor, Wout Joseph, "Performing QoE-measurements in an actual 3G network", 2010 IEEE International Symposium on Broadband Multimedia