



Performance evaluation of different techniques to estimate subjective quality in live video streaming applications over LTE-Advance mobile networks

Jaume Segura-Garcia, Santiago Felici-Castell^{*}, Miguel Garcia-Pineda^{**}

Departament de Informàtica, ETSE, Universitat de València, Avd. de la Universidad S/N, 46100, Burjassot, Valencia, Spain

ARTICLE INFO

Keywords:

Video quality assessment
Broadband wireless networks
Quality of experience
Video streaming
Factor analysis
Multinomial Linear Regression
Artificial neural networks

ABSTRACT

Current mobile service providers are offering Gigabit Internet access over LTE-Advanced networks. Traditional services, such as live video streaming, over wired networks are feasible on these networks. However different aspects should be taken into account due to the fast changing network conditions as well as the constrained resources of the mobile phones, in order to provide a good subjective video quality in terms of Mean Opinion Score (MOS). Our goal is to estimate and predict this subjective metric without information or reference from the original video, known as Non Reference approach. This approach is important for the Service Provider from a practical point of view, because it can keep the customer satisfaction at good levels. We analyze different estimation techniques running over a set of monitored variables throughout the whole steaming system, from the streaming server to the mobile phone. We have gathered variables related to bit stream, basic video quality metrics as well as Quality of Services variables. These variables are used to estimate MOS in a reliable and robust way. We compare three techniques such as Artificial Neural Networks (ANN), Factor Analysis (FA) and Multinomial Linear Regression, at different time scales and with Full Reference and Non Reference approaches. We carry out a performance evaluation of these techniques, concluding that the behavior of MOS estimation based on FA is more accurate, unless we had a lossless scenario related to Guaranteed Bit Rate services, where ANN performs better. The subjective video quality has been evaluated through surveys. Finally, we evaluate the accuracy of the estimated MOS against well known publicly available video quality algorithms following the recommendations given by Video Quality Experts Group (VQEG).

1. Introduction

Current mobile service providers are offering Gigabit Internet access over Long Term Evolution-Advanced (LTE-A) networks, specified by ITU-R 4G standards, called International Mobile Telecommunications Advanced (IMT-Advanced) (ITU-R, 2012a). These standards define services at 100 Mbps and 1 Gbps for high and low mobility communications, and the main difference with its predecessors is that these services are completely based on Internet Protocol (IP) and support Quality of Service (QoS). LTE-A technology (Ghosh et al., 2010) is standardized by 3GPP as IMT-Advanced compliant.

In this new framework, traditional services over wired networks are feasible, in particular those based in live video streaming (based on User Datagram Protocol (UDP) and Real Time Protocol (RTP)) such

as Internet Protocol Television (IPTV) and Video on Demand (VoD) (Oyman et al., 2010). However, different considerations should be taken into account due to the fast changing network conditions as well as the constrained resources of the mobile phones. In this case, Service Provider (SP) tries to maintain a good subjective quality of the delivered video, measured in terms of Mean Opinion Score (MOS) (ITU-R, 2012b) as a metric of the Quality of Experience (QoE) (Takahashi et al., 2008; Munyoz et al., 2013) in order to satisfy their subscribers. Nevertheless, subjective testing is time-consuming and requires special facilities.

In practice, subjective estimation is done using objective video quality metrics. But most of these metrics fail or are not sufficiently accurate, and/or require additional information that it is not available in a real scenario, such as the original video sequence. So, to define reliable and robust video quality metrics, we measure variables monitored

^{*} Corresponding author.

^{**} Corresponding author.

E-mail addresses: jsegura@uv.es (J. Segura-Garcia), felici@uv.es (S. Felici-Castell), migarpi@uv.es (M. Garcia-Pineda).

throughout the whole system, from the streaming server to the mobile phone or end user. Then, as many variables influence on QoE in these networks, our goal is to process these measured variables using three techniques, such as Factor Analysis (FA) (Gorsuch, 1983), Multinomial Linear Regression (MLR) and Artificial Neural Networks (ANN) based on Multi Layer Perceptron (MLP) (Rosenblatt, 1962) in order to estimate MOS (denoted by $\widehat{MOS}$), comparing their results with well-known objective video quality metrics. It must be stressed that we apply these techniques at different time scales and using two approaches, Full Reference (FR) and Non Reference (NR) (Chikkerur et al., 2011). FR requires information both the original (or reference) and the received video, while NR only use the information from the received video, that is more interesting from a practical point of view. Besides with NR, we consider two different approaches, the customer (called NR_{cu}) and the SP (called NR_{sp}) depending on the variables included. In order to generalize these estimations, we use different kinds of network services according to the different type of subscribers, based on different QoS requirements.

To evaluate the accuracy of the proposed metrics ($\widehat{MOS}$), we have measured the subjective video quality (in terms of MOS) with surveys, using standards ITU-R BT.500-13 (ITU-R, 2012b) and P.910 (ITU-R, 2008). In addition, we compare these metrics with well known publicly available video quality algorithms, following the recommendations given by Video Quality Experts Group (VQEG) (ITS, 2000).

The rest of the paper is structured as follows. Section 2 analyze the related work. Section 3 explains the network topology and test bench, defining the variables available in the different scenarios. Section 4 explains how we perform the subjective assessment of the received video quality. Section 5 describes the statistical methods and artificial neural networks used to estimate the objective video metrics. Section 6 details the results and shows the final expressions in order to define $\widehat{MOS}$ with different approaches. Section 7 analyzes and compares the performance of the proposed metric against well known video quality metrics. Finally, Section 8 concludes the paper.

2. Related work

Regarding objective video quality metrics, in the last decade have been proposed different FR and NR video quality metrics, such as MSAD, SSIM (Wang et al., 2004), 3SSIM (Li and Bovik, 2010), MSSSIM (Wang et al., 2003), STSSIM (Moorthy and Bovik, 2010), VQM (Pinson and Wolf, 2004) and O.23 (ITU-T, 1201), as shown in Table 1. It is must be stressed that O.23 metric (ITU-T P.1201 (ITU-T, 1201)) is NR, similar to our NR_{cu} approach, and it requires for its calculation: bandwidth, frame rate and size of I frames. A performance comparison of some of these FR video quality metrics are done in (Sedano et al., 2011) and (Moorthy et al., 2010) in a subjective manner using H.264/Advanced Video Coding (AVC), following the recommendations given by (ITS, 2000), concluding that the Spatial MOTion-based Video Integrity Evaluation (MOVIE (Seshadrinathan and Bovik, 2010)) index shows the high-

est performance and Temporal MOVIE the lowest among the studied metrics. In the study presented in (Chikkerur et al., 2011) concludes that MSSSIM, VQM and the perceptual spatio-temporal frequency-domain based on MOVIE indexes are the most reliable in terms of subjective MOS values. Besides for audio and video, the authors in (You et al., 2010) evaluate the different Full and Reduced Reference (which only requires partial information of the original video) quality metrics. They conclude that there are currently no objective metrics available that can replace subjective quality assessments and in particular, Peak Signal to Noise Ratio (PSNR) shows the worst results from the subjective point of view.

In relation to QoE estimation, an analysis of the different ITU standards and initiatives is shown in (Takahashi et al., 2008), to perform the objective quality assessments of IPTV services to estimate QoE in a subjective perspective. The different methods are categorized in media-layer models, parametric packet-layer models, parametric planning models, bit stream layer models and hybrid models. Our proposed methods are classified as this last one. In (Oyman et al., 2010) is analyzed both the potential and the limitations of 4G networks for delivering video content, from a theoretical point of view. The authors suggest that new performance evaluation methodologies should be done for these networks to account for various video quality metrics that involve human visual perception.

Besides, different techniques have been used to model QoE. In (Floris et al., 2012) the authors propose a QoE index for tablet devices based on a linear regression, taking into account only the bit rate, packet loss rate, play out delay and transmission interruption of H.264/AVC video sequences transmitted over lossy wireless channels. This index is only correlated with subjective MOS and is compared to SSIM and PSNR. In (Tsolkas et al., 2017), the authors provide a comprehensive guide to standardized and state-of-the-art quality assessment models. The authors identify and describe parametric QoE formulas for most popular service types (i.e., VoIP, online video, video streaming, etc.), indicating the Key Performance Indicators (KPIs) and major configuration parameters per type. In (Shafiq et al., 2014a), the authors model user engagement in video streaming applications on a real mobile network, as a QoE metric, with an 87% of accuracy. In particular, the authors gather different variables from the Radio Access Network (RAN) and Core Network (CN). From RAN they get information at the physical layer such as handovers, bitrate, signal strength, etc. and from CN they get flow-level information of video streaming traffic such as TCP flows information, flow duration, TCP flags, video traffic statistics (container type, encoding bit rate, video duration, etc.), etc. at network layer and higher. After a correlation analysis and applying machine learning, they conclude that using only CN variables, they get same accuracy to estimate user engagement than using both CN and RAN variables. In other words, due to this correlation and dependencies, information from higher layers above the physical layer, provides more reliable and better information. In (Leroux et al., 2011), the

Table 1
Objective video quality metrics.

Metric	Description
MSAD	Mean Absolute Difference of the color components in the correspondent points of each frame
SSIM	Structural SIMilarity is based on measuring of three components: luminance similarity, contrast similarity and structural similarity
3SSIM	3SSIM is based on region division (edges, textures and smooth regions) of source frames. Notice that the human eye can see differences on textured or edge regions more precisely than on smooth regions. The result metric is calculated as a weighted average of SSIM metric for those regions.
MSSSIM	MultiScale SSIM based on SSIM metric of several downscaled levels of original images. The result is a weighted average of those metrics. MSSSIM accounts for the multiscale nature of both natural images and human visual system
STSSIM	Spatio-Temporal video SSIM is based on the essence of MOVIE (Seshadrinathan and Bovik, 2010), with lower complexity. STSSIM uses motion information computed from a block-based motion-estimation algorithm and quality measures, using a localized set of oriented spatio-temporal filters
VQM	Video Quality Metric is DCT-based metric which exploits the property of visual perception and it is contained in ITU-T recommendation J.144 (ITU-R, 2001)
0.23 (P.1201)	Parametric non-intrusive MOS estimation for non-adaptive, progressive download type media streaming for lower resolution

Download English Version:

<https://daneshyari.com/en/article/6884804>

Download Persian Version:

<https://daneshyari.com/article/6884804>

[Daneshyari.com](https://daneshyari.com)