Quality of Experience Assessment in Internet TV

Joana Palhais[1], Rui S. Cruz[2], and Mário S. Nunes[2]

[1] Instituto Superior Técnico, Lisboa, Portugal
`joanapalhais@ist.utl.pt`
[2] Instituto Superior Técnico/INESC-ID/INOV, Lisboa, Portugal
`rui.cruz@ieee.org, mario.nunes@inov.pt`

Abstract. Nowadays, Service Providers are increasingly concerned about the concept of Quality of Experience (QoE), even more, when talking about Internet TV or WebTV, where no guarantees of delivery are provided. This paper describes the research and the results on the influence of the level of interest (on a particular sport) in the subjective quality assessment of the corresponding broadcasted media. This analysis is motivated by the work being developed in the European Project My-eDirector 2012, which has the capability to cover the London Olympic Games 2011 via the Web. Therefore, a subjective test was prepared and performed where each observer visioned and assessed the perceived video quality of a set of six sports, encoded in four different bitrate/resolution sets. From the analysis of the collected data it is possible to demonstrate that the interest level has a strong influence in the subjective assessment of the video quality. Based on these results, an empiric formula was deduced to estimate the Mean Opinion Source (MOS) as a function of bitrate and interest level.

Keywords: Mean Opinion Score, Objective Video Quality, Quality of Experience, Subjective Video Quality, Internet TV.

1 Introduction

The European Project My-eDirector 2012 [12] aims to develop an architecture for interactive and personalized WebTV. With this new architecture users will be able to choose the events they want to watch, the cameras that best capture the selected events or the athletes they want to follow. Due to the complexity of the architecture and the rich user interface of the terminal player, the evaluation by users of the media being displayed becomes difficult. It is thus necessary to develop specific assessment methodologies in order to define the QoE. For that purpose, a suite of tests with human evaluators needs to be performed to enable the collection of the corresponding subjective data, according to ITU-T Recommendation [5]. Such data must be validated, to obtain curves of MOS as a function of the evaluated parameters. With these results, an empirical additive formula must be deduced, to estimate the MOS as function of bitrate (R) and interest level (IL). This research is of surmount importance since, as far as

K. Pentikousis et al. (Eds.): MONAMI 2011, LNICST 97, pp. 261–274, 2012.

known by the authors, only Kortum & Sullivan [13] studied the influence that contents have on their subjective assessment, and all other works used generic movie clips, while My e-Director 2012 is focused on sport events.

The paper is organized as follows. After the Introduction, a review of the quality concepts Quality of Service (QoS), Quality of Perception (QoP) and QoE is exposed in Section 2. Section 3 identifies the differences between the subjective and objective methods, for video quality measurements, and the metrics that are more common in each category. Section 4 describes the process of test sessions preparation, as well as the implementation at the session day. Sections 5 and 6 present the results and the proposed formula for MOS estimation. Finally, Section 7 summarizes the conclusions that were obtained during the research and proposes future work to be done in the area.

2 Concepts Review

In the past, Service Providers were concerned about measuring the QoS for the audio and video data sent to their consumers. However, nowadays, more and more people are able to choose their own platform to watch video content. Regardless of the type of device, content viewed, or network used for access, each person still has some basic expectations about the viewing experience. This means that a new concept called QoE is rapidly growing up.

- **QoS and QoE:** are two distinct concepts that cannot be ignored and are both important. QoE is concerned with the overall experience that the user has when accessing and using the services, therefore, it is common to refer to QoE as a user-centered approach and to QoS as a technology-centered approach. QoS has been in use for a long time and has reached a high level of common understanding. ITU-T Recommendation E.800 [7] defines QoS as “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”. This concept is based on technical performance and typically measures the network performance at the packet level. The most common parameters used are packet loss, delay, jitter and throughput. The concept of QoE is relatively new and is attracting growing attention. Therefore, different definitions of the QoE are stated throughout the literature, as exposed in [10]. Despite of all these definitions, ITU-T Recommendation P.10/G.100 [6] defines QoE as “the overall acceptability of an application or service, as perceived subjectively by the end user”. This concept is based on the global enjoyment and satisfaction of the end user. Typically, the parameters more commonly used are fidelity of information, usability, responsiveness and availability.
- **QoP:** The concept of QoP emerged just after the QoE concept but focused in the detection of a change in quality or in the acceptance of a quality level. However, this is not a new concept as QoP was already known as the user-perceived QoS (QoSE). ITU-T [7] defines QoP or QoSE as “a statement expressing the level of quality that customers/users believe they have

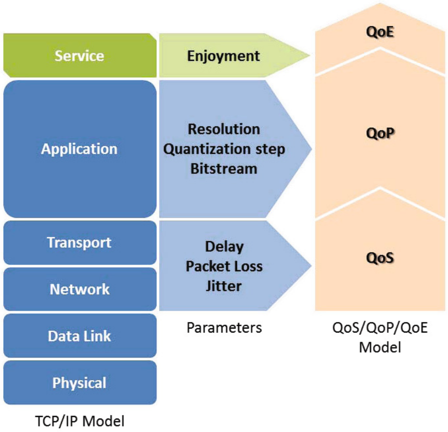


Fig. 1. Relationship between the TCP/IP model and the QoS/QoP/QoE three layers model

experienced". With the introduction of QoE, and in order to avoid ambiguities or exchange of concept meaning, it has been defined that the three quality concepts to consider would be QoS, QoP and QoE instead of QoS, QoSE and QoE. Typically, the QoP is measured with a subjective rating scale such as the MOS [2], a discrete numeric scale with values between 1 and 5, where 5 represents the highest quality and 1 the lowest. The major difference between QoP and QoE concepts is that QoP is specific for assessing video quality, while the QoE concept can be used to describe the evaluation of any type of experience, be it video or other typical daily routine (going to a restaurant, for example).

- **The co-existence of concepts:** QoS is a technical approach whereas QoE and QoP are user-centered approaches. The relationship between these concepts can be expressed in a three-layer model. QoS is the lowest layer (it operates at packet level) and QoE is the highest (related with the user opinion). This model can also be related to the TCP/IP model, as shown in Figure 1. It is intuitive that QoS should be the lower layer, QoP the intermediate and QoE the higher layer, due to their relationship with the TCP/IP layers. At the QoS level, the parameters used are those from the network and transport layers. These parameters help Service Providers to measure the network performance. To measure QoP and QoE levels, Service Providers must perform surveys over their clients to catch their perception on the quality of the service and the global satisfaction. By analyzing the results, Service Providers are able to know the maximum quality they can deliver and the sufficient level of quality that can be accepted by their viewers.

3 Video Quality Metrics

Video quality measurements are performed via objective and subjective methods. Objective methods use information contained in the image without the need of human observation. Subjective methods rely on the human judgment to infer the quality of the video. Regardless of the method used, results are usually reliable and correlated [1]. This section intends to clarify the differences between objective and subjective methods and to identify the metrics that are commonly used for the video quality assessment.

3.1 Objective Metrics

Objective video quality measurements do not need human intervention for classification of the video, as they are automated methods, based on algorithms, able to estimate the video quality by just analyzing the characteristics of the media stream. The metrics used are classified in three classes [10], being the first class the most common and accurate approach, and the last the less used:

- Full Reference (FR): both the original video and the decoded one are available;
- Reduced Reference (RR): some characteristics of the original video are used to compare with the decoded video;
- No Reference (NR): the original video is not available, only the decoded video.

Generally, Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) are the FR techniques used for objective metrics due to their simplicity, and both indicate the differences between the received video signal and the reference video signal. Other FR metrics can also be used, such as Perceptual Evaluation of Video Quality (PEVQ), Structural Similarity Index (SSIM) and Video Quality Metric (VQM). These latter metrics are more complex using not only the differences between frames, but also mechanisms to take into account the Human Visual System (HVS) and the perceptual effects of video impairments, in order to estimate how much a signal can be distorted until the human eye notices it.

3.2 Subjective Metrics

Subjective metrics are concerned about collecting data directly from end users and are recognized as the most reliable for quantifying the user's perception [9]. For that purpose, a group of observers must be recruited to obtain their opinion when asked to rate a sequence of videos or to detect a change in quality. The methodology followed for the tests is standardized in Recommendations ITU-R BT.500 [5] and ITU-T P.910 [8]. The advantage of this approach is that data is collected in a laboratory with a high level of control, by simulation of real environments through a controlled set of parameters such as, transmission delay or packet loss. The disadvantage is that the measurements are only concerned with the human ability to detect changes in quality, meaning that user's behaviors and interaction are not measured. There are several metrics that can be used such as, Double-Stimulus Impairment Scale (DSIS), Double-Stimulus Continuous Quality Scale (DSCQS), Single-Stimulus (SS), Single-Stimulus Continuous

Quality Evaluation (SSCQE) and Simultaneous Double-Stimulus for Continuous Evaluation (SDSCE). The differences between these metrics consist in showing or not a reference and in the type of the rating scale used.

4 Methodology for Subjective Assessment Tests

The methodology to prepare and setup the subjective tests session, covered materials and logistics, selection of observers and assessment rating scales. The subjective tests allow to infer the influence of content on the video quality assessment, and for that purpose, in each test session, a suite of six sports encoded in four different bitrates was prepared to be shown to each of the selected observers. At the end of each video clip the observers ranked their perceived video quality on a scale ranging from 0 to 10. During the sessions, the panel of observers experienced an environment similar to the one they are accustomed at their homes, where the original video is not available for comparison with the received decoded video.

4.1 Test Materials Selection

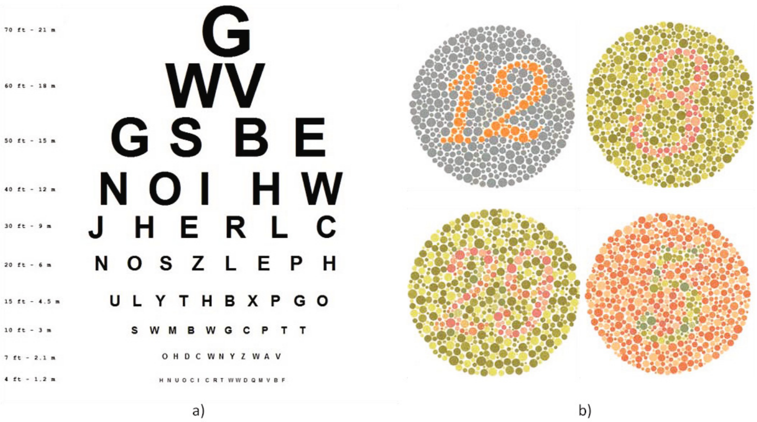
All of the test videos to be used were made public on the Internet (on a web server). The original videos selected for the tests are in high definition (HD) 720p format (resolution of 1280 x 720 pixels), coded with H.264 codec (in baseline profile) with 25 fps and a bitrate of 2 Mb/s. A total of thirty-two sport modalities were selected covering those typically viewed in the Olympic Games. In order to test the degradation of quality based on bitrate oscillation, the original videos were transcoded in four bitrate/resolution pairs (listed in Table 1) using the FFmpeg [3] tool functions. The resulting videos had no audio and were cut to a fixed duration of 30 s each.

4.2 Selection Criteria for the Observers

The test group should be formed by at least fifteen observers [5]. However, observers have to meet a certain set of prerequisites in order to be selected to participate, i.e., must be non-expert (not directly concerned or related with

Table 1. Bitrate/Resolution pairs for the transcoded videos

Bitrate (kb/s)	Resolution (pixels)
1450	848 x 480
600	424 x 240
350	320 x 176
190	320 x 176



a) b)

Fig. 2. Visual acuity tests: a) SnellenChart; b) Ishihara Plates

video quality as part of their normal work) and must be in their perfect visual conditions (pass in visual acuity tests), to ensure adequate video assessment test results. To test the visual acuity, including normal color perception, the candidates are subject to a simple vision test, at the test session day, by using the Snellen chart and the Ishihara plates, as depicted in Figure 2. If observers fail (do not pass either the Snellen or the Ishihara tests), they should not be accepted.

4.3 Subjective Assessment

During the video quality assessment tests, the observers rate the perceived video quality at the end of each video clip, i.e., at every 30 s, by selecting a rank value on a small window that pops-up over the client User Interface (UI) showing a star scale between 0 and 10. When the ranking window pops-up, the video sequence is paused to allow the observer to judge the viewed sequence and to rate it, avoiding therefore any type of pressure. The assessment star scale is a one-click scale and observers just need to choose the rating value by clicking over the respective star, after which the window automatically closes and the video sequence restarts for another 30 s clip. Each assessment process is cycled during the whole test session until the whole video sequence is watched and rated. This type of subjective analysis is a NR method, since the observer only has access to the decoded video. The method used for these tests is a trade-off between the SS and SSCQE metrics, as these metrics only require the decoded video, despite of their different rating scales. However, in the SSCQE, videos must correspond to paired programme segment (PS) and quality parameter (QP). The option was therefore to use a discrete scale, such as the one used by SS, with the videos arranged with a PS of sports and a QP of bitrate. For the subjective test, these PS/QP pairs are watched in a pseudo-random order.

4.4 Architecture for the Assessment Tests

The architecture for the assessment session is very simple and is composed by a web streaming server (accessible from the Internet and the intranet of the test facilities) and N networked client computers in the test room with access to the intranet and Internet. Note that, the network of the test facilities does not cause any type of impairments in the test, i.e., no delays or playback errors. The web streaming server stores all the available video files used for the tests and provides a database (DB) to register all the ratings given by each observer via the respective client UI. Clients use a web browser (with a special media player) to request media streams from the server. There will be as many client computers as the number of observers for the tests to be performed individually. The database built to store the classifications of the video clips, is pre-populated with the ID of each observer, the list of six sports that each observer will watch and the order of visualization of the videos. The list with the six sports contains the three sports that each observer likes more and the three that he/she likes less (the same methodology described in [13]). It is expected that observers tolerate more errors (greater image degradation) in their favorite sports and give lower scores to sports less desired (a similar behavior observed on YouTube followers, that do not mind watching a poorer quality video, if it fulfills their needs, than not having access to it). To display/assess the videos, a browser based media player was developed using Microsoft Silverlight framework [11]. The media browser needs a log in access by introducing the observer's ID. This allows the application to identify the database record with the assessment video sequence for that specific observer. After log in, the observer enters in the viewing environment, a media player that shows the videos in full screen mode and has no trick-function buttons for interaction. The interaction with the application occurs only at the end of each video, when a window pops-up with the assessment scale. The log out is done automatically by the media player, as soon as the observer classifies the last video clip in the sequence.

4.5 At Session Day

An initial online survey was previously sent to a wide group of students from the two *campi* of a University, asking about name, gender, age, profession, e-mail address and a rank, from 1 to 5, on the interest level for each of the thirty-two sports listed. With this information it was possible to identify the video quality experts, the list of the six sports that best fit the users' interests, and pre-select the candidates for the test. There were performed two test sessions in different rooms, the ambient light was constantly monitored by using a light-meter to maintain an average ambience luminance of 200 lux [4]. Prior to each test session, the observers were tested for their visual acuity, receiving their ID if approved, and subsequently introduced to the methodology of video quality assessment, instructed on the grading scale, the actions to take, the duration of the test session and prepared with a training sequence of four videos to clarify any doubts that might arise. Each assessment session lasted fifteen minutes during which the observers were only concerned on the video quality assessment they rated at the end of each video sequence. The twenty-four video clips were shown sequentially, but with short intervals between each one to allow the observer to assess the content just watched. The four rate/resolution pairs were shown in a random way to avoid ranking by impulse.

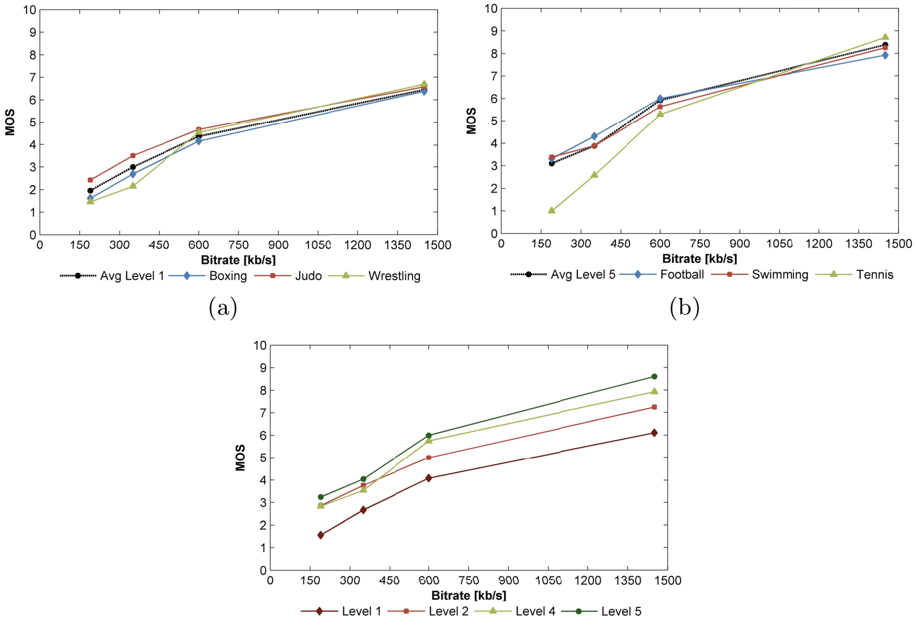
5 Data Analysis

From the initial survey, 268 responses were collected and 260 validated, from which, 24 were selected for the subjective tests. A total of 144 videos were therefore viewed, grouped by interest level (IL) from 1 to 5. Due to lack of sufficient data, IL 3 was not considered in this analysis. The sports corresponding to low interest levels were Boxing, Wrestling and Judo (33% of preferences), and those corresponding to high interest levels were Football, Swimming and Tennis (35% of preferences). Computing the average for each IL (regardless of the sport modality) produced the results plotted in Figure 3(c), turning evident that the observers tend to value more (around two values scale points higher) a video with the same bitrate, just because they have higher interest on it. Comparing side by side the MOS of the three most watched sports, it is evident in Figure 3(a) that for low interest, at the highest bitrate (1450 kb/s) all have a ranking almost coincident with the average for IL 1. For 600 kb/s, the ratings are also close to the average value, but for the two lowest bitrates, the ratings are more dispersed with more errors in the estimation, despite being within the limits, reinforcing the hypothesis that the observer tolerance has a high influence in the results. The same comparison for sports more watched with high interest level, Figure 3(b), makes also evident that the rankings given to Football and Swimming are very similar between them for all bitrates and are in agreement with the average result. However, although Tennis ratings are within the limits for the highest bitrate, for the other bitrates, 190 kb/s and 350 kb/s, the ratings are outside the limits. This phenomenon can be explained by the high movement that characterizes Tennis as the players are constantly running from one side to the other of the court and the ball is very small, reaching very high speeds. These results indicate that the interest level positively influences the rank. As the interest level increases, the ranking also increases. Another interesting conclusion is that, except for the lowest interest level, at 190 kb/s and 350 kb/s, observers almost do not notice differences in quality, since the rating assigned to those is less than 1 scale value.

6 MOS Estimation

With the collected data it is possible to express the MOS as a function of the bitrate R and the interest level IL. The goal is to establish an additive formula, where the first term depends only on R and the second on IL, expecting an equation in the form of 1:

$$MOS = f_1(R) + f_2(IL) \tag{1}$$



(a) (b)

(c)

Fig. 3. Comparison of MOS between levels: (a) for low interest sports; (b) for high interest sports; (c) by interest level and bitrate

By drawing the trend lines for each interest level, it is possible to estimate the equation for $f_1(R)$, and so, the equations for each trend line become:

$$\text{Level 1}: \qquad y = 5.2264 \log_{10}(x) - 10.458 \tag{2}$$
$$\text{Level 2}: \qquad y = 5.0374 \log_{10}(x) - 8.8389 \tag{3}$$
$$\text{Level 4}: \qquad y = 6.0535 \log_{10}(x) - 11.277 \tag{4}$$
$$\text{Level 5}: \qquad y = 6.2987 \log_{10}(x) - 11.477 \tag{5}$$

Observing these equations for each IL, it appears that they are quite similar, especially between the two lowest and the two highest levels. It can then be inferred that it is possible to obtain a function of R which approximates the behavior of each interest level. Averaging the trend lines of those expressions, the first term for the general MOS function 6 is achieved:

$$f_1(R) = 5.6540 \log_{10}(R) - 10.513 \tag{6}$$

Equation 6 describes the MOS as a function of bitrate, regardless of the interest level. Therefore, the second term, that depends on IL, is clearly used to level the MOS. Studying the MOS as function of interest levels for each available bitrate the graph of Figure 4 is obtained. This graph shows that MOS also has a logarithmic behavior for each interest level, pointing out to a second term, also with a logarithmic behavior, despite being less pronounced in the two lowest levels. As the resulting curve should provide a good approximation for the average

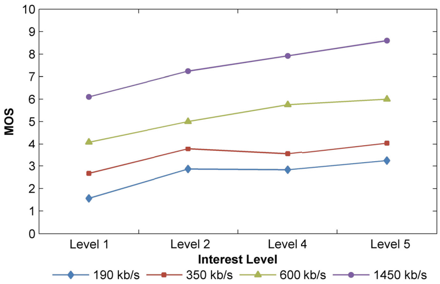


Fig. 4. MOS as function of interest level for each available bitrate

curves obtained for each interest level, it is possible then to estimate $f_2(IL)$, by keeping in mind that it is also a logarithmic function, resulting therefore in 7:

$$f_2(IL) = 2.6318 \log_{10}(IL) - 1.041 \tag{7}$$

Combining the terms, the empiric MOS formulation can then be expressed as 8:

$$MOS = 5.6540 \log_{10}(R) + 2.6318 \log_{10}(IL) - 11.554 \tag{8}$$

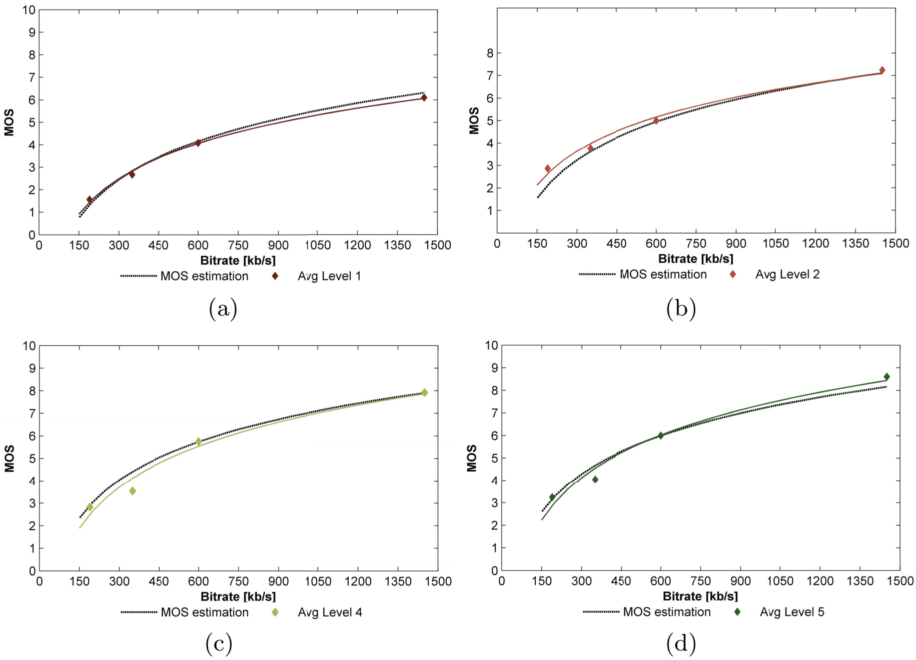
Figure 5: a) to d) shows the MOS computation for each interest level using 8, where it is possible to verify that the new MOS formulation provides a really good estimate for each interest level, with a standard deviation between 0.10 and 0.12. With formulation 1 only, the standard deviation would be 0.5, but expressing the MOS as a function of R and IL it comes reduced more than four times. Despite these quite good results, there is still the need to introduce a new parameter in the estimation of MOS, related with sports with high temporal activity (TA), such as Tennis. For this sport the average curve does not represents a good approximation for the two lowest bitrates, forcing us to conclude that a new parameter will be required, as function of TA.

6.1 Temporal Activity

The temporal activity can be estimated by computing the difference, pixel by pixel, between two successive frames. ITU-T Recommendation [6] defines temporal activity (TA) as the “maximum value of standard deviation found along the video frames”, as expressed in 9:

$$TA = \max\{std\ [F_n(i,j) - F_{n-1}(i,j)]\} \tag{9}$$

In this equation, $F_n(i,j)$ is the pixel at the ith row and jth column of nth frame in time. However, for sequences with changes of camera, the resulting



(a)

(b)

(c)

(d)

Fig. 5. MOS as function of IL and available bitrate: (a) Level 1; (b) Level 2; (c) Level 4; (d) Level5

temporal activity can have a high value even if the video has a low temporal activity. In sports capturing, many changes of camera may occur, and in order to minimize and smooth this effect, the 99% percentile should be applied to the global temporal activity[1] The main problem, now, relies on the identification of temporal activity characteristic for each sport. Grouping sports at a high level, turns possible to establish the following three (TA) levels:

- low temporal activity: $TA < 35$;
- medium temporal activity: $35 < TA < 50$;
- high temporal activity: $TA > 50$.

With this approach, Tennis can be identified as a high temporal activity sport, confirming the experimental verification. Javelin can also be considered a high temporal activity sport, since for the lowest bitrates the javelin cannot be identified in the air, confirming again that the temporal activity stages provides a good characterization. However, for BMX, Diving, Pommel Horse and Taekwondo, the high temporal activity stage does not apparently match, since the experimental results do not reveal such behavior. Although these sports are typically slow movement sports with one or two athletes, it is common to capture the event with several changes of camera. The cameraman is always looking for new plans,

---

[1] The 99% percentile values were obtained using MATLAB®.

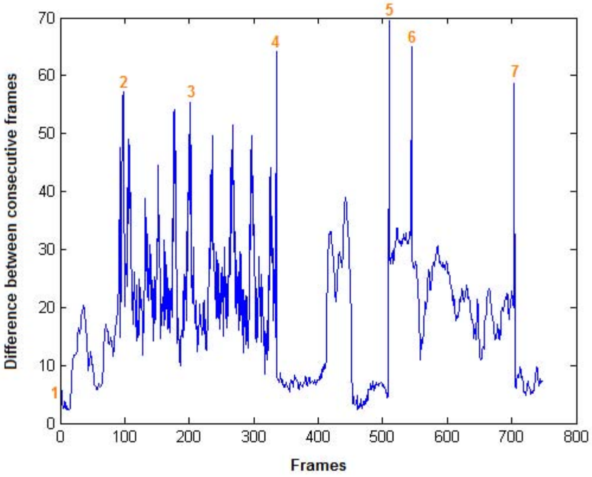


Fig. 6. Difference between consecutive frames to Pommel Horse

making zoom-ins and zoom-outs. Due to this non-intrinsic behavior of the sport, the difference between consecutive frames can be significant even when the 99% percentile is taken into account. Figure 6 shows the graph obtained for Pommel Horse, which is full of peaks. However, only seven of these peaks really represent camera changes. The other peaks are due to the camera movement to follow the exercise along the pommel horse. The peaks representing changes of camera are identified in the graph, with tags numbered 1 to 7. Figure 6 shows that despite Pommel Horse having only six explicit changes of camera, the 99% percentile does not eliminates all the existing intermediate peaks, which are due to the camera movements in following the athlete, therefore classifying Pommel Horse with high TA. The same reasoning is valid to BMX, Diving and Taekwondo. Due to this phenomenon, the 99% percentile cannot smooth the effect of these peaks, but using the 95% percentile, it is possible to reduce the intermediate peaks effect, preventing these sports to be classified at the high TA stage, when they intrinsically should not. For sports with low and medium TA the empiric formula 8 can be used, but for sports with high TA a new empiric formula (10) should be used:

$$MOS = \begin{cases} 5.6540\log_{10}(R) + 2.6318\log_{10}(IL) - 11.554, & TA < 50 \\ \text{new equation to develop} & TA > 50 \end{cases} \tag{10}$$

Due to the lack of data, the new formula for sports with high TA could not be developed in due time for this paper, but left for future work in this field.

7 Conclusion and Future Work

The results obtained, allowed concluding that the interest level has a positive influence on the subjective rating as, for the same content, observers tend to increase the ratings (for the same bitrate) only because they feel more interested. Between the lowest and the highest IL, the difference in MOS can reach 2.5 values and this result is independent of the type of sport. However, for sports of high interest and high temporal activity, the difference raises up to 2 values below the average, while for sports of low interest at low bitrates, the difference is around 1 value from the average. It is possible therefore to conclude that the developed empiric MOS formula, as a function of R and IL, provides a good approximation for the MOS on almost all the sports, but still has to take into account another parameter related with TA, in order to have a more general application. However, since only two sports were identified in the high temporal activity stage, Tennis and Javelin, it makes no sense to introduce at this stage the TA parameter in the MOS expression, essentially due to the lack of data collected for these two sports. For that purpose, the best solution would be to develop another formula related to sports with high TA and integrate it in the general MOS formula. Additional research still needs to be done in this area, with a larger and more diversified group of observers, in order to collect data with statistical relevance to allow tuning the parameters for all dimensions, but essentially for the temporal activity parameter, namely:

- Sports with highTA, such as Tennis and Javelin: The performed test only had enough data to evaluate the Tennis behavior to IL 5. Other interest levels and sports must be analyzed to verify if the same phenomenon can be clearly identified.
- Sports with a medium IL: Due to lack of sufficient data, IL 3 was not considered, since only one observer has watched one sport with this interest level.
- Test more bitrates between the 190 kb/s and 1450 kb/s, to obtain smoother curves: The performed tests only considered four bitrates, with a gap of information between the 600 kb/s and the 1450 kb/s.

Acknowledgments. The research leading to these results has received funding from the European Union's Seventh Framework Programme ([FP7/2007-2013]) under grant agreement ICT-215248.

References

[1] Dialogic. Quality of Experience for Mobile Video Users. White Paper 11681-01, Dialogic Corporation (December 2009), `http://www.dialogic.com/medialabs/`

[2] ETSI. Human Factors (HF); Quality of Experience (QoE) requirements for real-time communication services. Technical Report TR 102 643, European Telecommunications Standards Institute (January 2010)

[3] FFmpeg. FFmpeg - the Open Source Multimedia System (2011), `http://www.ffmpeg.org/`

[4] ITU-R. Methodology for the subjective assessment of the quality of television pictures. Recommendation BT.500-11, International Telecommunication Union - Radiocommunication Sector (2002)

[5] ITU-R. Methodology for the subjective assessment of the quality of television pictures. Recommendation BT.500-12, International Telecommunication Union - Radiocommunication Sector (September 2009)

[6] ITU-T. Vocabulary for performance and quality of service, Amendment 2: New definitions for inclusion in Recommendation ITU-T P.10/G.100. Recommendation P.10/G.100 (07/2008), International Telecommunication Union - Telecommunication Standardization Sector (July 2007)

[7] ITU-T. Definitions of terms related to quality of service. Recommendation E.800, International Telecommunication Union - Telecommunication Standardization Sector (2008)

[8] ITU-T. Subjective video quality assessment methods for multimedia applications. Recommendation P.910, International Telecommunication Union - Telecommunication Standardization Sector (2008)

[9] Kim, H.J., Lee, D.H., Lee, J.M., Lee, K.H., Lyu, W., Choi, S.G.: The QoE Evaluation Method through the QoS-QoE Correlation Model. In: Proceedings of the Fourth International Conference on Networked Computing and Advanced Information Management, NCM 2008, vol. 2, pp. 719–725 (September 2008), doi:10.1109/NCM.2008.202

[10] Zapater, M.N., Bressan, G.: A Proposed Approach for Quality of Experience Assurance of IPTV. In: Proceedings of the First International Conference on the Digital Society (ICDS 2007), p. 25. IEEE (2007), doi:10.1109/ICDS.2007.4

[11] Microsoft Corporation. Microsoft Media Platform: Player Framework v2.5 (formerly Silverlight Media Framework) (2011), `http://smf.codeplex.com/`

[12] My eDirector. My eDirector 2012 Website (2009), `http://www.myedirector2012.eu`

[13] Kortum, P., Sullivan, M.: The Effect of Content Desirability on Subjective Video Quality Ratings. Human Factors 52(1), 105–118 (2010), doi:10.1177/0018720810366020