STUDYING THE CHALLENGES IN ASSESSING THE PERCEIVED QUALITY OF MOBILE-PHONE BASED VIDEO

Selim Ickin[1], Lucjan Janowski[2], Katarzyna Wac[3], Markus Fiedler[1]

Blekinge Institute of Technology[1], AGH University of Science and Technology[2], University of Geneva[3]

ABSTRACT

Evaluating video Quality of Experience (QoE) on a mobile phone has not yet been studied much. It is common that the data collected through user studies in mobile platform involves high fluctuation of user ratings without obvious reasons related to variation in network level. *User disparity*, user's various intermediate or previous experiences, video bitrate, and the objective measure of criticality are a few of the reasons that need to be identified with a well-designed user experiment. We present an experiment procedure to understand better the perceived quality of video in mobile platform. First, we investigate the reliability of the data, and identify *unreliable* users. Then, we investigate the psychological influence factors of previous experiences on the recent perceived quality known as the *memory effect*, and the influence of the bitrate on the time it takes for users to react and evaluate the video quality, *i.e.*, *user response time*.

Index Terms— Quality of Experience, user response time, video bitrate, video content, memory effect

1. INTRODUCTION

The growing popularity of video-based applications used for video sharing opens new interesting research topics in Quality of Experience (QoE) research in mobile platforms. For the welfare of long term organisation success, service providers need improved QoE models to predict the user satisfaction particularly on high resource demanding applications like video streaming. There have been efforts in the literature to find trends between frame rate, bit rate, display size, video content [1], and the QoE, which is referred to the most commonly used evaluation method for measuring perceived quality, *Mean Opinion Score (MOS)* [2]. However, the majority of QoE models in the literature do not consider much the temporal dynamics and historical experiences of user's satisfaction with application or service performance [3]. As John Locke has pointed out back in 17th century [4], *"lukewarm water can feel hot or cold depending on the hand touching it was previously in hot or cold water"*, the previous and the recent experiences play an important role on

the instantaneous judgment of a particular video quality. In addition, there may exist large disparity between users, which makes QoE user studies a challenge. It's hard to understand "why" and "when" the users react to a given visual stimuli in a mobile phone. Fully understanding the fluctuation in the ratings of each user and also rely on those ratings is a difficult task. On the other hand, identifying unexpected fluctuation of users' MOS ratings with a proper experiment design enables more accurate QoE models. In this paper, we present an experiment procedure to have a better understanding in the assessment of the video QoE in mobile platform. Throughout our study, we concentrate on three items; user disparity, memory effect, and user's response time. First, we investigate the disparity of the user ratings among all participants in the study. Next, we observe the changes of user's rating for exactly the same video, investigate the memory effect and the influence of video bitrate on the user perception based on the analysis of the user disparity. Lastly, the influence of the video bitrate on the user response time is studied.

This paper is organized as follows. Related work is discussed in Section 2. The detailed description of the measurement setup and the method are given in Section 3. Section 4 presents the results, and it is concluded with final remarks in Section 5. The limitations and future work within this user study are given in Section 6.

2. RELATED WORK

The importance of diversity in user ratings is stated in [5] and a Standard Deviation of Opinion Scores (SOS) parameter is introduced to test the reliability of the conducted user studies, which can be examined throughout the analysis and interpretation of the QoE results. The diversity within a single individual user is studied in [6] and it's claimed that at any point in time, individuals can have different conflicting views. [7] proposes a method for user modeling that distinguishes users based on how products are perceived by different users. *"Framing effect"* is a term commonly used in psychology which states that the decision of a particular option by the user is dependent on how that option is presented to the user [8]. It is stated in [9] that what is enjoyably learnt is less likely to be forgotten. Asymmetrical traces are observed in user perception such that, users imidiately respond

negatively to degradations in quality while being calmer in rewarding the increased quality [10]. In [11], five major drawbacks of MOS scale are claimed such as its reflection in human memory, recency effect, variation in the perception for the same evidences with time, variation of interpretation of the words in MOS scale among users, and the necessity for huge number of subjects. In [12], primacy and recency effects are emphasized and discussed; elements at the beginning or end of the sequence have increased weigthings in the evaluation of the presented information. It is claimed that the duration of impairment has little impact on quality ratings but strong impact by the peak intensity experienced during the video. In addition, recency effects are reduced when continuous assessment was used. The Interdependancy of QoE and QoS (IQX) hypothesis, proposed and verified for packet loss in [13], presents an exponential functional relationship between QoE and Quality of Service (QoS). In [14], a lightweight kernel module called MintMOS is introduced in order to infer real time video streaming and MOS values are presented with respect to three varying parameters; encoding bitrate, delay and packet loss rate. The varying durations between the time of stalling events and the timing of the user ratings, as an evidence to the user reaction, are observed in previous studies [15]. It is observed that the users rate low MOS values immediately after the stalling events.

3. METHOD AND EXPERIMENTAL DESIGN

The experiments were carried out with 30 participants, which are randomly chosen among students at Blekinge Institute of Technology (BTH) in Sweden. Among 30 participants, 80% were male and the age range was 21 to 30 years.
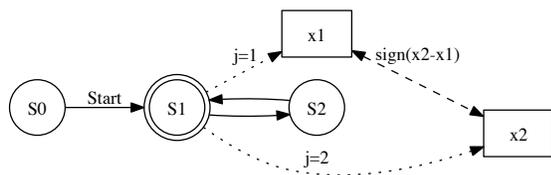
Fig. 1. The state diagram illustrates the procedure followed during the experiments.

Table 1. Complementary table that describes the experiment design and the states involved for each scenario.

$S1$ \ $S2$	100 kbit/s	300 kbit/s	900 kbit/s
100 kbit/s	-	100 kbit/s(1)	100 kbit/s(2)
300 kbit/s	300 kbit/s(3)	-	300 kbit/s(4)
900 kbit/s	900 kbit/s(5)	900 kbit/s(6)	-

We did not set any restrictions for participants regarding the proximity of the phone to the user during the experiments. The users were allowed to take flexible and comfortable position, *i.e.*, the position they are accustomed to take while using their private smart phone in daily life. Identical instructions were given to all users prior to the experiments with a special training session, where they watch three different movies of different qualities from three different video sources. We have used a set of videos which are freely available online [16]. The videos were presented in full screen landscape mode in 480×800 pixels.

We have set up a QoE measurement environment in which the video clips were presented in predefined order through Samsung Galaxy S with super active-matrix organic light-emitting diode (AMOLED) capacitive touchscreen with 16 million colors. The experimental design consists of 18 different videos, which each are compressed with bitrates; $100\,\text{kbit/s}$, $300\,\text{kbit/s}$ and $900\,\text{kbit/s}$. In [17], it's claimed that the difference between the preceding stimuli should be adequate enough in order for the user to perceive the difference between the recent and the previous quality. The choice of bitrates are chosen with a factor of 3, according to the Weber-Fechner law. After each video play-out, the user was requested to rate the video quality where the choices for ratings were vertically oriented buttons marked within 5 point MOS scale [2] as a response to the question: *"Overall, rate the experience for the previous sequence"*. The rating menu pops up at the end of each sequence play-out. A clean background with a neutral gray [18] is shown to the user between the play-out of each sequence. The reason for this is to imitate flushing of user memory to some extent. Practically, each experiment conducted on the mobile phone lasted approximately 10 minutes per participant.

The experiment consists of six different deterministic scenarios and each scenario consists of three deterministic states $S0$ (training session), $S1$ (session when users are asked to watch and rate a particular video clip with certain quality), and $S2$ (session when users are asked to watch the same videos that are presented at $S1$ with different quality) as presented in Fig. 1. The content of the video that was shown to the users at the training session, *i.e.*, $S0$, was different than the ones in the experiment. After $S2$, users are asked to watch and rate exactly the same video that has been watched in $S1$, again. We differentiate the first and second visits to state $S1$ by introducing a visit count variable, $j \in \{1, 2\}$. The MOS ratings that are received by the users are denoted with $x \in \{1, 2, 3, 4, 5\}$. x_1 and x_2 are recorded in the first and second visit to $S1$, respectively. Each state and scenario consists of three different video clips, and the group of video clips that are used for each scenario are different. By this way, we aimed to imitate flushing of user memory, and to filter out the possible biases, *e.g.*, video content dependency of user perception.

3.1. User Disparity

Not every user's perception in the study is identical, *e.g.*, each participant has his/her own unique perception thus follows different MOS rating patterns for the same given video sequences. Identifying the rating patterns for each user with the same given experiment design helps QoE evaluation to trace users whose ratings are statistically different from the majority. The linear Pearson correlation coefficient per bitrate for one user versus all users is calculated in [19], and the users whose correlations are less than 0.75 are excluded from the dataset. Based on our results, we classified the users whose correlations are greater than 0.7 as *"Correlated"* users. Then, we have counted the number of unexpected ratings per each user, and classified the users as *"Expected"*, whose unexpected ratings are lower than 2. A simple example for an unexpected rating is a decrease in rating with respect to the increase of bit rate of the same consecutive video. We consider these traces as indication when participants were not fully focussed to the experiments. We continued with further analysis on the data such as, memory effect and user response time, in two approaches; one with *"All"* user data, and the latter with *"Filtered"* data, *i.e.*, without the users whose ratings do not satisfy the *correlated* and *expected* criteria.

3.2. Memory Effect

The study of autocorrelation of perceived instantaneous quality shows that the perceived instantaneous quality is influenced by the latest received quality [20]. We investigated this through the presentation with visual stimuli, *i.e.*, repeated video sequences, to the participants and studied the fluctuation of MOS values on the same video of a particular user and the memory effect to understand the psychological influence factor of previous video sequences. We look for the answer to the question: *"Does the order of the sequences with different qualities, e.g., bitrates, affect the user's perceived quality of the video?"*. The first, second and the fourth scenarios of the experiment are designed so that the first and the last bitrates of the videos are identical with intermediate video of relatively higher bitrate. Similarly, in the third, fifth, and the sixth scenarios, the bitrate of the videos of the initial and the last video are identical with intermediate video of relatively lower bitrate. The function that calculates the difference in between the ratings, $\text{sign}(x_2\text{-}x_1)$, is calculated for each scenario. The video sources with 100 kbit/s bitrates are presented to the users twice with an intermediate video with 300 kbit/s in scenario 1 and with an intermediate video with 900 kbit/s in scenario 2. The video sources with 300 kbit/s bitrates are presented to the users twice with an intermediate video with 100 kbit/s in scenario 3 and with an intermediate video with 900 kbit/s in scenario 4. The video sources with 900 kbit/s bitrates are presented to the users twice with an intermediate video with 100 kbit/s in scenario 5 and with an intermediate video with 300 kbit/s in scenario 6. An additional table that

describes the experimental procedure is presented in Table 1, where the first column is $S1$, and the first row is $S2$ as depicted with the bitrates presented at each state. The matching column-row pair represents video bitrates presented at the second visit to $S1$ with the corresponding scenario number shown in parenthesis. The analysis are applied with two different ways; first with *"All"* users, second with *"Filtered"* users that bear only *correlated* and *expected* MOS values.

3.3. User Response Time (T_R)

The reasons for varying user response times with respect to the visual stimuli should be identified. T_R is calculated as the time difference between popping up of the rating menu and the reception of the MOS value from the user. We want to understand whether there exists any influence of video bitrate on the user's thinking time, *i.e.*, T_R, for a particular video quality. Therefore, we analysed T_R with three different levels of bitrates; 100 kbit/s, 300 kbit/s, and 900 kbit/s.

3.4. Metric of Criticality (o)

In order to analyse the above mentioned three items, it's important to understand the motion complexities of the video sequences that are used throughout the experiments and to understand the similarities among them. Therefore, we have calculated the objective metric, o, for all video sequences. Temporal Information (TI) measures frame-to-frame image changes and it is computed based on two consecutive frames [21] and Spatial Information (SI) measures the spatial detail as given in Eq. 1 and Eq. 2. TI and SI values are used in the calculation of the objective measure of criticality, o, in Eq. 3 [22]. In the below equations, $F(n)$ denotes luminance channel of the nth video frame, *Sobel* represents the filter type, and rmsspace is the root mean square over an entire video frame. The o values are calculated for all the video sequences used in the study.

$$TI(n) = \text{RMS}_{\text{space}}[F(n) - F(n-1)] \quad (1)$$

$$SI(n) = \text{RMS}_{\text{space}}[Sobel(F(n))] \quad (2)$$

$$o = \log_{10}(\text{mean}_n[SI(n) \cdot TI(n)]) \quad (3)$$

4. RESULTS

In total, the experiments lasted 355 minutes with the participation of 30 users. 54 ratings were collected from each participant that makes 1620 ratings in total.

One reaction from the participants was *"I have seen the same video, but I don't remember what I rated previously, so my ratings might be different"*. This was related to the memory factor that the users might not remember how exactly they have perceived the previous identical video. This causes

the users to re-evaluate the same video with lately gained experiences, rather than pushing the same button again. Feedbacks regarding the MOS scale was also observed, one phrase was, *"My perception for a few videos were between the scale $3-$fair and $4-$good, and I was unsure about the rating"*. Another interesting feedback was *"The finger-prints remaining on the ratings affects my perceived quality of the video for the next sequences"*. This was considered as a relevant feedback for any future QoE studies conducted in touch screen devices.

4.1. User Disparity

We investigated some users with different MOS rating patterns than the majority when the same stimuli is applied within the same experiment design. Users 6, 8, 9, 12, 16, 19, 25, 27, and 28 bear ratings with correlations less than 0.7. Among these nine users, 6, 9, and 16 bear also unexpected ratings, *i.e.*, the number of unexpected ratings on the data for those users are 5. Therefore, we performed a deeper analysis on those three users to identify anomalies in the ratings. The video source (k), the order of two consecutive videos, criticality (o), and the MOS values are given for those 3 users in Table 2. The o values for those sources with the corresponding bit rates are also studied to understand the differences of the objective qualities. The close values of o metric can be considered as the cause for the unexpected ratings, however, in general, by looking at Table 2, this is not the case. In fact, we observe an increase in o and the bit rate

Table 2. A detailed look at the users whose ratings by the users whose ratings are not classified as *correlated* and not *expected*.

User	Subject 6									
Source(k)	2		10		18		8		7	
Order	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Bitrate(kbit/s)	100	300	100	900	300	100	300	100	300	100
Criticality(o)	6.95	7.03	7.59	7.73	5.92	6.15	8.18	8.08	8.53	8.50
MOS	2	1	3	2	1	2	1	2	2	3
User	Subject 9									
Source(k)	14		23		9		24		25	
Order	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Bitrate(kbit/s)	300	900	300	900	300	900	900	100	900	300
Criticality(o)	7.27	7.29	7.26	7.32	8.53	8.57	6.19	6.15	5.06	5.06
MOS	5	4	5	3	5	4	2	3	3	4
User	Subject 16									
Source(k)	2		18		14		23		21	
Order	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Bitrate(kbit/s)	100	300	300	100	300	900	300	900	900	100
Criticality(o)	6.95	7.03	6.05	5.92	7.27	7.29	7.26	7.32	6.35	6.31
MOS	5	4	3	4	5	4	2	3	3	4

but a decrease in the MOS values, and vice versa. Therefore, we conclude that users 6, 9, and 16 bear uncorrelated and unexpected ratings and classified them as unreliable users. We have continued our analysis with two different data sets; a first one("All" data set) with all 1620 instances and the latter with 1458 ratings, by filtering out those 3 users("Filtered" data set).

4.2. Memory Effect

The psychological influence factor, namely the memory effect, on the obtained models is analysed in this section. The difference between the scenarios, *i.e.*, the particular order of presentation of video clips of various contents and bitrates, play an important role on the MOS values. This percentages of increases and decreases in ratings are given in Table 3 by considering both *"All"*, and *"Filtered"* data sets. It is observed from Table 3 that the ratings remain the same with 51% or below for all scenarios. This fluctuation of the ratings can be well explained by the influence of the quality with the lately watched video clips. Although there are unchanged ratings for the identical videos, there are also considerable amount of traces, which shows that the first and the last ratings, *i.e.*, x_1 and x_2, for exactly the same video are different. For all scenarios except for scenario 5, the first and second ratings, *e.g.*, x_1 and x_2 are statistically significantly $(p = 0.05)$ different. In scenario 5, the videos of the best quality is repeated at $S1$ with a gap of 60 seconds in between. It can be interpreted that the extremely good qualities are remembered better despite the intermediate varying qualities, and that yields consistent user ratings. We have not observed significant changes on this analysis when three users are filtered out as discussed before.

Independent of the detailed analysis of the scenarios, the mean MOS values versus the bitrates (k) are also calculated. Similarly, we suggest logarithmic-law models as in Eq. 4 (R^2=0.96) and in Eq. 5 (R^2=0.98) for *"All"* and *"Filtered"* data sets, respectively. As expected, when unreliable users are filtered from the data set, there is an increase in the number of decreased MOS ratings in scenarios one and two,

Table 3. The percentages of increases, decreases and number of changes of the user MOS ratings for each scenario.

MOS	Increase(%)		Decrease(%)		NoChange(%)	
Scenario	All	Filtered	All	Filtered	All	Filtered
1	36	36	14	15	50	49
2	16	14	41	44	43	42
3	43	44	12	10	45	46
4	19	19	40	39	41	42
5	30	30	27	26	43	44
6	13	15	36	35	51	50
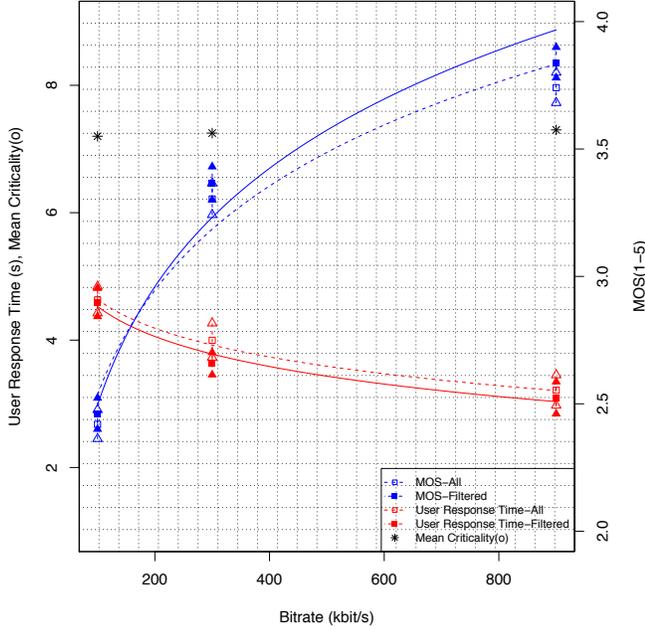
Fig. 2. T_R, o, and MOS vs video bitrate are illustrated with and without the identified unreliable users with suggested logarithmic-law models (CI=95 %).

and an increase in the increased ratings in scenarios three and six. Therefore we see an improved model without the unreliable users. We need to note that these models should not be considered as a ground truth, but only show us clues about the behaviour of the model, since the bitrates we have considered are limited to only three levels, *e.g.*, 100 kbit/s, 300 kbit/s, and 900 kbit/s.

$$MOS_{All} = 0.59 \ln k - 0.18 \qquad (4)$$

$$MOS_{Filtered} = 0.67 \ln k - 0.59 \qquad (5)$$

o values for all the sources with three different bit rates are calculated and the mean o is illustrated on the left axis with the corresponding bitrate in Fig. 2. Analysis with more data and to obtain more robust models is scheduled for future work.

4.3. User Response Time (T_R)

We studied the influence of video bitrate on T_R, and analysed the user data first with *"All"*, then with *"Filtered"* data set.

The T_R values with respect to the video bitrates are illustrated in Fig. 2. Logarithmic-law models can be suggested for both versions of the data. The dashed line illustrates *"All"* data set with the suggested fit Eq. 6 (R^2=0.99), while the continuous line with Eq. 7 (R^2=0.97) illustrates the *"Filtered"* data set.

$$T_R(s)_{All} = -0.65 \ln k + 7.63 \qquad (6)$$

$$T_R(s)_{Filtered} = -0.68 \ln k + 7.66 \qquad (7)$$

By looking at the analysis, we can conclude that the User Response Time (T_R) for a particular video is influenced by its quality, and it decreases with the higher video bitrate. The reason, could be that the higher bitrates with good quality is perceived as good without any doubts, meanwhile the ones with lower bitrates need more time for the users to think and give the final decision.

5. CONCLUSION

Challenges in the assessment of video QoE on the mobile platform such as user disparity, memory effect, and User Response Time (T_R) are investigated throughout six detailed scenarios that consist of repeating video sequences of different quality. Prior to the analysis of the user data, we have identified the correlations of ratings with the bitrates and also spotted the suspicious ratings for each user. 3 out of 30 users are confirmed as unreliable. Then, we continued our analysis with two different data sets to reckon the differences in results, with and without those identified unreliable users. Next, the number of fluctuated ratings with respect to the same video clip are investigated and it has been observed that the ratings received with respect to a particular video clip had tendency to change statistically in 5 out of 6 scenarios when the same particular video clip is presented to the user for the second time after 60 seconds, given that they have watched the same video clip with different quality within this duration. The influence of the video bitrate to the user perception is studied as well as its influence on the T_R. A logarithmic-law model between the MOS and the video bitrate is suggested. Also, a logarithmic-law model between the video bitrate and the response time is recommended that shows lower user response times with higher video bitrates. One reason for this could be that users perceive and interpret the high qualities easier as compared to the low quality ones. Foreseeing the discussed challenges during the assessment of video QoE would help to build more robust QoE models and more precise MOS predictions.

6. FUTURE WORK

We have presented our suggested models in relation to MOS, User Response Time (T_R), and the video bitrate, however these suggested models can be improved with consideration of more data points for future work. The bandwidth demand is higher in high motion sequences than the low motion sequences, which in reality would make users rate a high motion video quality differently, knowing that the video is high network demanding. During the analysis of the memory effect, the focus is on the detection of deviations in the ratings of a particular user during the experiments, therefore the user's previous experiences, which might have taken as basis for the

judgments, are neglected. It is important to analyze in future, all the factors influencing the perceived quality of the video in mobile platforms and to derive more robust models for QoE.

7. REFERENCES

[1] J. Joskowicz and J. C. L. Ardao, "Combining the effects of frame rate, bit rate, display size and video content in a parametric video quality model," in *Proceedings of the 6th Latin America Networking Conference*, New York, NY, USA, 2011, LANC '11, pp. 4–11, ACM.

[2] Transmission Quality, "Methods for subjective determination of transmission quality," *ITU Recommendation P800*, vol. 800, pp. 800–899, 1996.

[3] T. Hosfeld, S. Biedermann, R. Schatz, A. Platzer, S. Egger, and M. Fiedler, "The memory effect and its implications on web QoE modeling," in *Teletraffic Congress (ITC), 2011 23rd International*, sept. 2011, pp. 103 – 110.

[4] "Contrast effect," `http://en.wikipedia.org/wiki/Contrast_effect`, 2012.

[5] T. Hobfeld, R. Schatz, and S. Egger, "SOS: The MOS is not enough!," in *Quality of Multimedia Experience (QoMEX), 2011 Third International Workshop on*, sept. 2011, pp. 131 –136.

[6] E. Karapanos, J. Martens, and M. Hassenzahl, "Accounting for diversity in subjective judgments," in *Proceedings of the 27th international conference on Human factors in computing systems*, New York, NY, USA, 2009, CHI '09, pp. 639–648, ACM.

[7] E. Karapanos and J. Martens, "Characterizing the diversity in users' perceptions," in *Proceedings of the 11th IFIP TC 13 international conference on Human-computer interaction*, Berlin, Heidelberg, 2007, INTERACT'07, pp. 515–518, Springer-Verlag.

[8] J. N. Druckman, "Evaluating framing effects," *Journal of Economic Psychology 22*, 2001.

[9] C. Ardito, M. F. Costabile, R. Lanzilotti, and F. Montinaro, "Towards the evaluation of UX," Towards a UX Manifesto, COST294-MAUSE affiliated workshop, 2007.

[10] N. Cranley, P. Perry, and L. Murphy, "User perception of adapting video quality," *Int. J. Hum.-Comput. Stud.*, vol. 64, pp. 637–647, August 2006.

[11] C. Kuan-Ta, T. Cheng-Chun, and X. Wei-Cheng, "Oneclick: A framework for measuring network quality of experience," in *INFOCOM 2009, IEEE*, april 2009, pp. 702 –710.

[12] D. S. Hands and S. E. Avons, "Recency and duration neglect in subjective assessment of television picture quality," *Applied Cognitive Psychology*, vol. 15, no. 6, pp. 639–657, 2001.

[13] T. Hossfeld, P. Tran-Gia, and M. Fiedler, "Quantification of quality of experience for edge-based applications," in *Proceedings of the 20th international teletraffic conference on Managing traffic performance in converged networks*, Berlin, Heidelberg, 2007, ITC20'07, pp. 361–373, Springer-Verlag.

[14] M. Venkataraman and M. Chatterjee, "Inferring video QoE in real time," *Network, IEEE*, vol. 25, no. 1, pp. 4 –13, january-february 2011.

[15] S. Ickin, K. De Vogeleer, M. Fiedler, and D. Erman, "The effects of packet delay variation on the perceptual quality of video," in *Local Computer Networks (LCN), 2010 IEEE 35th Conference on*, oct. 2010, pp. 663 –668.

[16] "Video quality experts group," `http://www.its.bldrdoc.gov/vqeg/vqeg-home.aspx`, 2012.

[17] P. Reichl, S. Egger, R. Schatz, and A. D'Alconzo, "The logarithmic nature of QoE and the role of the weber-fechner law in QoE assessment," in *Communications (ICC), 2010 IEEE International Conference on*, may 2010, pp. 1 –5.

[18] C. Fenimore and et. al., "Perceptual effects of noise in digital video compression," *140th SMPTE Technical Conference*, 1998.

[19] VQEG HDTV Group, "Test plan for evaluation of video quality models for use with high definition tv content," `http://www.its.bldrdoc.gov/media/5871/vqeg_hdtv_testplan_v3_1.doc`, 2009.

[20] F. Guyard and S. Beker, "Towards real-time anomalies monitoring for QoE indicators," *Annales des Telecommunications*, pp. 59–71, 2010.

[21] L. Janowski, P. Romaniak, and Z. Papir, "Content driven QoE assessment for video frame rate and frame resolution reduction," *Multimedia Tools and Applications*, pp. 1–18, 2011.

[22] P. Ramos, F. Vidal, and Raquel Prez Leal, "Perceived Video Quality Estimation from Spatial and Temporal Information Contents and Network Performance Parameters in IPTV," in *Proceedings of the 2010 Fifth International Conference on Digital Telecommunications*, Washington, DC, USA, 2010, ICDT '10, pp. 128–131, IEEE Computer Society.