ASSESSING QUALITY OF EXPERIENCE FOR ADAPTIVE HTTP VIDEO STREAMING

Jingteng Xue[1], *Dong-Qing Zhang*[2], *Heather Yu*[2], *Chang Wen Chen*[1]

University at Buffalo, SUNY[1] Huawei Technologies USA[2]
{jingteng,chencw}@buffalo.edu {dongqing.zhang,heatheryu}@huawei.com

ABSTRACT

In this paper, we present a novel subjective quality model for online adaptive movie streaming service. The proposed model considers the Quality of Experience (QoE) of streaming video viewing as a cumulative evaluation process of consecutive segments that compose a story line. Under bandwidth constraint, streaming client may select lower-quality segment, pause playback for re-buffering, or both. The momentary QoE loss at these events are weighted by the content suspension level. Meanwhile such experience penalty remains to influence user's opinion for the rest of the service. If the picture becomes too noisy or the interruptions occur consistently, the user may stop watching. The proposed scheme includes two parts. First, a parametric model estimates the quality loss of a interfered segment based on its network-level packet characteristics. Second, a cumulative function integrates the impact of streaming events. Both steps demand minimum computation and can be updated in real time. A subjective test to train and validate the proposed parametric model is designed and performed. This model is fundamentally different from all existing QoE assessment schemes in that temporally cumulative viewing experience of the users, instead of simple global statistics, is evaluated.

Index Terms— QoE, adaptive HTTP video streaming, DASH, HLS, instantaneous quality, cumulative quality, subjective test, parametric model

1. INTRODUCTION

Currently video streaming over Hyper-Text Transfer Protocol (HTTP) significantly out-populates conventional UDP based streaming protocols due to the advantages of simplicity and robustness in HTTP. Adaptive HTTP streaming protocols such as HTTP Live Streaming (HLS) and Dynamic Adaptive Streaming over HTTP (DASH) achieve decoder driven rate adaptation by providing multiple versions of bit-streams in a variety of bit-rates and breaking these bit-streams into sequences of small HTTP file segments. Throughout the streaming process, the receiver adaptively switches among the available bit-rates by selecting the corresponding bit-stream segments to match the playback rate and the instantaneous TCP throughput.

Adaptive video streaming system requires a quality measurement mechanism. Not only that service providers would like to evaluate the service quality but also the resource allocation as well as the adaptive bit-stream switching algorithm demand a quality orientated guideline. Quality of Service (QoS) captures objective and system-related characteristics but does not align well with subjective quality in the complexity of adaptive video streaming, because user experience is closely related to both picture fidelity and playback smoothness. While better image quality demands larger size segments, the increased video bit-rate may lead to playback interruptions under bandwidth constraint. The eventual perceived quality is a combinational result of the two factors. It is crucial to develop Quality of Experience (QoE) models that better match true human experiences.

Although the current standards of HTTP streaming only unify the control message for QoE reporting without specifying any methods of quality measurement [1], in practice the performance of any streaming system relies heavily on the effectiveness of the adopted QoE model to enhance its ultimate service performance.

1.1. Related work

General quality assessment of static video playback has been throughly studied. Various approaches including the successful schemes of SSIM [2] and MOVIE [3] have been developed with their strategies ranging from full reference to no reference. They work very well in offline video playback scenarios where the playback procedure can be accurately predicted. Streaming experience, on the other hand, will be significantly influenced by the random fluctuation of bandwidth. The adaptive playback process needs to be determined dynamically in real time. Research in QoE modeling for adaptive HTTP streaming is still at its early stage.

In [4], the QoE of HTTP streaming is defined to be the probability of decoder re-buffering, which has been widely used in many related studies. However, this over-simplified model is not a good indicator of the pattern of playback jitter or the image quality. The image quality indicator is especially important under the mechanism of bit-stream switching

This research is supported by Gift Funding from Huawei Technologies, Inc.

in adaptive HTTP streaming.

Recently there are some works trying to characterize the temporal stalling factor. The freeze distortion model [5] measures the quality loss of playback interruptions without considering the image quality. This does not fit well with adaptive HTTP streaming characterized by quality selection. The QoE metric reported in [6] considers the pattern of jitters and local content importance by subjective training of the content. This approach demonstrates the significance of temporal factor in HTTP streaming QoE. However this is essentially a human-in-the-loop process that is virtually impossible to be generalized for regular streaming applications. QDASH [7] can approximate the QoE of bit-stream transition processes. But it does not provide a close form model to monitor and report system QoE, which is necessary in many adaptive HTTP streaming systems. PSQA [8] tries to solve this problem by training a random neural network to simulate the way human reacts to video streaming. The resulting neural network is then used to generate a subjective-like quality report in the streaming assessments. Although the idea is to capture human vision characteristics, the model itself works like a black box that does not explicitly map specific playback impairments in the overall quality loss.

One common problem among all these existing approaches is the lack of a simple yet effective way to characterize the inherent connections between the two important factors of streaming in both content quality and playback smoothness. The parametric model we develop in this research is intended to work towards tackling this major challenge in the design of a QoE model for adaptive HTTP streaming system.

1.2. Proposed scheme and major contribution

We consider the QoE of adaptive streaming video a cumulative experience driven by streaming events including interruption and quality switching. The instantaneous perceptual quality is characterized by the segment coding parameters. The quality loss caused by interruption is weighted by the intensity of the local storyline, which is estimated by the segment packet properties. The proposed model generates a comprehensive QoE by pooling all instantaneous experience.

The proposed pooling function characterizes temporal exponential decay, which is a fundamental property of human memory retention [9]. Such property has been studied and utilized in the field of Web QoE [10] as well as video quality. The ITU standard [11] suggests that human memory effects can be modeled as a decaying weighting function.

The contributions of this research are in three folds.

First, we formulate an adaptive video streaming QoE model to account for the exponential decay of human memory. This research is so far the first endeavor that recognizes and utilizes the forgetting curve property of human memory in the study of HTTP streaming QoE. The proposed study not only connects the two factors between content quality and playback smoothness, but also reveals the relationship between the streaming events and the cumulated experience.

Second, on top of the proposed QoE model, we derive a parametric approach that estimates the real-time instantaneous and cumulative QoE. We illustrate the principle using one specific metric in this paper. However we believe that virtually any other video quality assessment (VQA) scheme can be substituted into the proposed framework for an optimized tradeoff between performance and complexity.

Third, we design and conduct a subjective test with thirty participants and twenty simulated streaming profiles that covers a wide variety of streaming scenarios. The dataset is used to train and validate the proposed model. The simulation results demonstrate desired performance in prediction.

2. ADAPTIVE STREAMING QoE MODEL

2.1. Instantaneous experience of playback

The distortion of a rendered picture $\mathbf{f}'_i$ in the pixel domain can be estimated by its MSE against the original picture $\mathbf{f}_i$

$$d_i = \mathbb{E}\left\|\mathbf{f}_i - \mathbf{f}'_i\right\|^2. \tag{1}$$

Consider a common coding structure of IPPP, where each reconstructed inter-predicted frame $\mathbf{f}'_{i,coded}$ is the sum of the previously rendered picture $\mathbf{f'}_{i-1,coded}$ and a quantized residual signal $\mathbf{R'}_i$,

$$\begin{aligned} \mathbf{f'}_{i,coded} &= \mathbf{f'}_{i-1,coded} + \mathbf{R'}_i, \\ \mathbf{f}_i &= \mathbf{f'}_{i-1,coded} + \mathbf{R}_i. \end{aligned} \tag{2}$$

$\mathbf{R'}_i$ is an approximation of the real differential signal $\mathbf{R}_i$ between the original frame and the reference decoded frame. If the target picture is delivered and reconstructed in this way, the observed distortion will be:

$$d_{i,playback} = \mathbb{E}\left\|\mathbf{f}_i - \mathbf{f}'_{i,coded}\right\|^2 = \mathbb{E}\left\|\mathbf{R}_i - \mathbf{R}'_i\right\|^2. \tag{3}$$

Such distortion is due to lossy compression of the residual signal and can be fully determined by the quantization parameter (QP).

We use a linear model to predict quality from QP [12] in order to estimate the picture quality $q_{k,coded}$ with QP_k and set the model parameters a and b to bound $q_k \in (0,1]$ for all preset QP values. We then define it as the instantaneous quality x of a coded frame

$$x_{k,playback} = q_{k,coded} = a\,\mathrm{QP}_k + b. \tag{4}$$

2.2. Instantaneous experience of interruption

We estimate the perceived quality of an interruption moment to be the opposite of the last seen picture quality weighted by the intensity of the interrupted scene.

Consider an interruption occurs at i when picture f'_{i-1} freezes. The instantaneous perception distortion $d_{i,stall}$ is the information loss:

$$\begin{aligned} \mathbf{f}'_{i,stall} &= \mathbf{f}'_{i-1,playback}, \\ d_{i,stall} &= \mathbb{E}\left\|\mathbf{f}_i - \mathbf{f}'_{i,stall}\right\|^2 = \mathbb{E}\left\|\mathbf{R}_i\right\|^2. \end{aligned} \tag{5}$$

Here $d_{i,stall}$ indicates the intensity of the broken scene. The fluctuation of the temporal complexity of video content is reflected by the variance of bit-rate change for the compressed frames in the bit-stream. Given a fixed QP, a frame consisting of more bits is considered to contain richer motion information and represents a more complex scene. The motion complexity generally correlates with the intensity of the scene.

Through simulations, it is observed that the encoder compression ratio is approximately constant within a scene, with the exception of IDR frames which should be ignored in our model. For a stalling moment $k \in [i, j]$, j the last stalling moment, we use $r_{i-1,coded}$ the bit count of the last decoded and displayed frame to linearly approximate the residual norm of frame k.

$$\mathbb{E}\left\|\tilde{\mathbf{R}}_k\right\|^2 = \frac{r_j^2}{\eta}. \tag{6}$$

To further characterize the logarithmic relation in rate distortion theory we apply a logarithmic operation in the weight function

$$W_{i,stall}^{unbound} = log(\frac{r_{i-1,coded}^2}{\eta}) \tag{7}$$

In order to bound W in $(0, 1]$, we consider r_{QP} the upper bound of frame size (or segment size) with the specified QP. We use r_{QP} to linearly scale W. From (7) we have

$$W_{i,stall} = \frac{\log\big(\min(r_{i-1,coded}, r_{\mathrm{QP}})\big) + c}{\log(r_{\mathrm{QP}}) + c}, \tag{8}$$

where $c = -log(\eta)/2$ is a model constant.

In order to estimate r_{QP}, we assume that the size of frames with a fixed QP is governed by a Laplace distribution. This assumption is reasonable given the independent identically distributed property of compressed frames.

Given the distribution, we can find the effective upper bound r_{QP} by keeping $P(r > r_{\mathrm{QP}}) \leqslant \epsilon$, ϵ a small threshold. Given the mean μ_{QP} and the variance ν_{QP}, we have

$$P(r > r_{\mathrm{QP}}) = \frac{1}{2} e^{-\frac{r_{\mathrm{QP}} - \mu_{\mathrm{QP}}}{\sqrt{\nu_{\mathrm{QP}}/2}}} \leqslant \epsilon \tag{9}$$

$$r_{\mathrm{QP}} \geqslant \mu_r - \sqrt{\frac{\nu_r}{2}} \log(2\epsilon) \tag{10}$$

The distribution parameters μ_{QP} and ν_{QP} of the frame sizes with the specified QP for a certain bit-stream can be retrieved via an online algorithm.

The instantaneous quality of any frame in this interruption period $[i, j]$ will be:

$$x_{k,stall} = -W_{i,stall}\, x_{i-1} \tag{11}$$

2.3. Instantaneous quality of initial buffering

In order to predict the QoE of initial buffering, we define an initial QP that approximately reflects the average quality of the streaming service. We also assume initial expectation is a constant. In this way, the initial buffering time is proportional to the cumulated experience loss. The initial values are: $W_{init} = 0.5$, $x_{init} = a\mathrm{QP}_{init} + b$, $\mathrm{QP}_{init} = 27$.

$$x_0 = -W_{init}\, x_{init} \tag{12}$$

2.4. Cumulative quality

We have introduced in previous section an immediate quality function $x(t)$ that provides the *instantaneous experience* at moment t during initial buffering, playback, or interruption. To characterize the cumulative experience in viewing streaming video, we need to weight this instantaneous experience by the forgetting curve of human perception starting from the beginning of the playback until the time k of the measurement. For the convenience of derivation, we will start with a continuous instantaneous experience function. Once the cumulative function is derived, we shall convert this continuous function into a discrete filter for easy QoE measurement.

We first define the cumulative experience as an integration of instantaneous experience over all time until the moment of measurement. This can be written as:

$$Q(t) = \int_{-\infty}^{t} x(\tau) e^{-\xi(t-\tau)}\, \mathrm{d}\tau. \tag{13}$$

Consider a short time period of ι from α to β (i.e. $\beta - \alpha = \iota$) in which $x(t) \equiv x_\beta$, we have

$$\begin{aligned} Q(\beta) &= \int_{-\infty}^{\beta} x(\tau) e^{-\xi(\beta-\tau)}\, \mathrm{d}\tau \\ &= e^{-\xi\iota} Q(\alpha) + (1 - e^{-\xi\iota}) x_\beta / \xi. \end{aligned} \tag{14}$$

For the application of video streaming, the basic time unit is the playback time of a single frame, which equals $1/f_{fr}$, f_{fr} being the frame rate. We can re-write (14) in its discrete form as:

$$y_k = \gamma\, y_{k-1} + (1 - \gamma)\, x_k, \tag{15}$$

where k represents a time slice, x_k is the scaled instantaneous quality at time k, $\gamma = e^{-\xi/f_{fr}}$ characterizes the memory strength, and y is the cumulative quality. Equation (15) is essentially an IIR filter.

In adaptive HTTP streaming systems, a video bit-stream is segmented into small chunks. We assume all frames within one segment share the same encoding configuration with same QP value. A typical segmenter like FFMPEG Open Source Segmenter splits video with fixed time duration t_{seg}. Therefore, each segment approximately contains a constant number of frames. For variable bit-rate (VBR) video, the bit-rate of a segment is proportional to its file size. In addition, an HTTP streaming player generally holds a basic time unit t_{rebuf} for network caused playback interruption in order to avoid jitter.

Consider the time slice τ which equals to the greatest common divisor of t_{seg} and t_{rebuf}. Within a slice $\big((p-1)\tau, p\tau\big]$ there are $k = \tau f_{\text{fr}}$ frames with the same instantaneous experience $x_{p\tau}$. From (15), we will have

$$\begin{aligned} y_{p\tau} &= \gamma^k y_{p(\tau-1)} + \sum_{i=0}^{k-1} (1-\gamma)\gamma^i \, x_{p\tau} \\ &= \gamma^k y_{p(\tau-1)} + (1-\gamma^k) x_{p\tau}, \end{aligned} \tag{16}$$

which provides a macro version of (15) for the HTTP streaming when the video is segmented into chunks of multiple frames. In the following text, we use γ to represent γ^k in (16).

For this study, we use the same memory strength parameter γ to model the memory effect in the initial buffering stage and in the playback interruption events since initial buffering can be seen as a special case of the playback interruption.

In summary, we have derived in this section a parametric model for instantaneous experience . An accumulation as in (16) generates accumulated experience for all moment in the viewing process. In this model, memory strength γ is unknown and needs to be estimated. We shall design and perform a subjective test to provide the training dataset for the validation of the proposed model.

3. SUBJECTIVE TEST

We designed and performed a subjective test in order to train and evaluate the proposed QoE model. In the design, we address the three most common quality impairments of adaptive HTTP streaming: (i) initial buffering time; (ii) quality of video segments (and the switching between them); and (iii) interruptions due to re-buffering.

3.1. Streaming profiles and test sequences

We manually generate twenty streaming profiles. Each profile describes a series of events in a HTTP adaptive streaming, including initial buffering time, selected segments and re-buffering time. Together, these profiles simulate a wide variety of streaming scenarios.

Table 1 lists the profile configurations. Each row represents a profile with the sequence name, the segmentation length (seg len) in seconds, and the streaming event history. The numbers in the "Timeslot" columns are segment ID. All segments have the same length in time as indicated by "seg len". An empty slot indicates stalling. Each segment is coded into two qualities. A "'" attached to an ID number (e.g. 1', 2', . . .) indicates lower quality.

Table 1. Configuration of the simulated streaming profiles

id	Sequence	Seg Len	Timeslot (k)						
			0	1	2	3	4	5	6
1	Prince	3			0	1	2	3	4
2	Prince	3		0'	1'	2'	3'	4'	
3	Killer	3	0			1	2	3	4
4	Killer	3	0		1'	2'	3'	4'	
5	HotTub	3	0	1			2	3	4
6	HotTub	3	0	1		2'	3'	4'	
7	Football	2	0	1	2			3	4
8	Football	2	0	1	2		3'	4'	
9	Soccer	2	0	1	2	3			4
10	Soccer	2	0	1	2	3		4'	
11	Touchdown	3		0	1	2	3	4	
12	Touchdown	3	0'	1'	2'	3'	4'		
13	Crowd	2	0		1	2	3	4	
14	Crowd	2	0	1'	2'	3'	4'		
15	Ducks	2	0	1		2	3	4	
16	Ducks	2	0	1	2'	3'	4'		
17	Factory	3	0	1	2		3	4	
18	Factory	3	0	1	2	3'	4'		
19	Dinner	3	0	1	2	3		4	
20	Dinner	3	0	1	2	3	4'		

Seven standard video test sequences and three movie trailers have been used in this test, each with two profiles. Following the recommendations of the subjective test methodologies, the test sequences are ten to fifteen seconds. The tailored clips are segmented with time slots equal to "seg len" and compressed by x264 into QP 27 and QP 37 quality levels. The bitstreams are then encapsulated by MP4Box to generate the .mp4 files. In total there are 100 video segments.

3.2. Test methodology

The subjective tests are performed via Internet browser. We build a website "dashSimulator" that is able to strictly replicate an adaptive streaming process given its profiled event script. The initial buffering and the playback interruptions are simulated via controlled delays.

The tests follow SS (Single Stimulus) test methodology [11]. The viewer is given the presentations one by one. After each presentation, the viewer is asked to rate the streaming quality and the content via web controllers. We use a scale of zero to ten to allow viewers to catch subtle quality difference.

Thirty assessors participated in this subjective test. They all have normal eyesight. Seven of them are experts in video technology. Despite the outliers, there are a total of 470 valid data points collected. The ratings are scaled into $[0, 1]$ and averaged among the viewers.

The collected subjective test results are depicted in Figure 1. This box plot marks the accepted mean opinion score

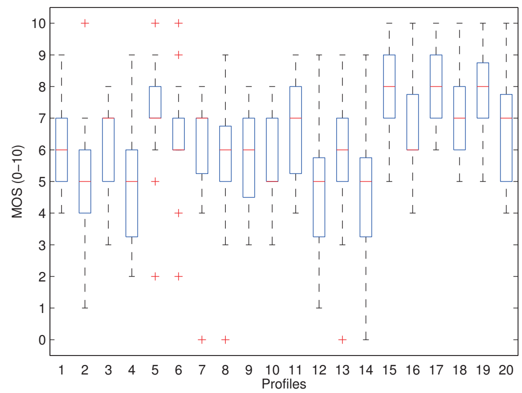

Fig. 1. Collected subjective test results

(MOS) range of each streaming profile. The red crosses indicate outliers which were excluded. The MOS are then calculated.

4. TRAINING AND EVALUATION

In (16), the memory strength γ is unknown and needs to be determined. In this section, we derive the optimal γ with a subjective test dataset. We then verify the proposed model by applying the derived γ to estimate the QoE of the tests in the dataset and comparing the results with the ground truth.

4.1. Training methodology

Consider a streaming process u in time period $1 \dots k$ with observed eventual QoE $\hat{y}_k^{(u)}$. The sequence of playback events in process u is represented by vector $\mathbf{x}^{(u)} = \left[x_1^{(u)} x_2^{(u)} \dots x_k^{(u)} 1\right]$.

As discussed in Section 2, different values of quality indicator x reflect not only different bit-stream quality (caused by different QP of 27 and 37) but also the stalling history.

Expand (15) to get $\tilde{y}_k^{(u)}$, the estimated QoE of the proposed model becomes

$$\begin{aligned}\tilde{y}_k^{(u)} &= \sum_{i=1}^{k} \gamma^{k-i}(1-\gamma)x_i^{(u)} + \gamma^k y_0^{(u)} \\ &= \mathbf{x}^{(u)} \left[\gamma^{k-1}(1-\gamma), \dots, \gamma^0(1-\gamma), \gamma^k y_0^{(u)}\right]^T\end{aligned} \tag{17}$$

A series of u tests has identical initial status $y_0^{(1)} = y_0^{(2)} = \cdots = y_0^{(u)}$ and known results $\hat{\mathbf{y}} = \left[\begin{array}{cccc} y^{(1)} & y^{(2)} & \dots & y^{(u)} \end{array}\right]^{\mathrm{T}}$. We estimate the QoE with a linear model

$$\tilde{\mathbf{y}} = \left[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(u)}\right]^T [z_0, z_1, \dots, z_k]^T \triangleq \mathbf{X}\,\mathbf{Z}. \tag{18}$$

Table 2. Results of 4-fold cross validation

	Group1	Group2	Group3	Group 4
γ	0.67	0.69	0.71	0.64
PCC	0.79	0.83	0.85	0.81

First, we need to find the optimized $\mathbf{Z}$ that minimizes the least square error

$$\begin{aligned}\mathbf{J} &= \|\hat{\mathbf{y}} - \tilde{\mathbf{y}}\|^2 = (\hat{\mathbf{y}} - \tilde{\mathbf{y}})^{\mathbf{T}}(\hat{\mathbf{y}} - \tilde{\mathbf{y}}) \\ &= \hat{\mathbf{y}}^{\mathbf{T}}\hat{\mathbf{y}} - \hat{\mathbf{y}}^{\mathbf{T}}\tilde{\mathbf{y}} - \tilde{\mathbf{y}}^{\mathbf{T}}\hat{\mathbf{y}} + \tilde{\mathbf{y}}^{\mathbf{T}}\tilde{\mathbf{y}} \\ &= \hat{\mathbf{y}}^{\mathbf{T}}\hat{\mathbf{y}} - \hat{\mathbf{y}}^{\mathbf{T}}\mathbf{XZ} - \mathbf{Z}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\hat{\mathbf{y}} + \mathbf{Z}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\mathbf{XZ}\end{aligned} \tag{19}$$

Set the partial derivative of (19) to be zero, we have

$$\begin{aligned}\frac{\partial}{\partial \mathbf{Z}^{\mathbf{T}}}\mathbf{J} &= -\mathbf{X}^{\mathbf{T}}\hat{\mathbf{y}} - \mathbf{X}^{\mathbf{T}}\hat{\mathbf{y}} + \mathbf{X}^{\mathbf{T}}\mathbf{XZ} + \mathbf{X}^{\mathbf{T}}\mathbf{XZ} \\ &= -\mathbf{2X}^{\mathbf{T}}\hat{\mathbf{y}} + \mathbf{2X}^{\mathbf{T}}\mathbf{XZ} = \vec{0}\end{aligned} \tag{20}$$

Therefore,

$$\mathbf{Z} = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-\mathbf{1}}\mathbf{X}^{\mathbf{T}}\hat{\mathbf{y}} \tag{21}$$

Since the quality indicator $\mathbf{x}$ is known. We can calculate $\mathbf{Z}$.

Second, we estimate the optimized γ using the $\mathbf{Z}$ obtained above. According to (17), $\mathbf{Z}$ provides k samples of γ, $\gamma_1 = z_0/z_1$, $\gamma_2 = z_1/z_2$, …, $\gamma_k = z_{k-1}/z_k$. An optimized γ can then be derived by

$$\gamma = \frac{1}{k}\sum_{i=1}^{k}\frac{z_{i-1}}{z_i} \tag{22}$$

4.2. Cross validation

Four fold cross validation is performed. The twenty data points are randomly partitioned into four groups. In each group, the unselected fifteen data points are used for training purpose. The training follows (21) and (22) to output a γ. The trained γ will then be used in the parametric model to predict the selected five data points. The Pearson correlation between the prediction set and the MOS set is calculated to measure the performance. The results are shown in Table 2.

From the results we can see that the variation of either γ or PCC from the four groups is very small. This fact demonstrates the stability of the proposed parametric model.

Second it is shown that the proposed parametric model predicts well the subjective experience on the dataset.

4.3. Evaluation results and discussion

We picked the best γ from Group3 to run on the entire dataset. The resulting correlation between the prediction and the subjective test scores are shown in Figure 2.

We also implemented the re-buffering based quality model [4] on the dataset (shown in Figure 2 as blue dots).

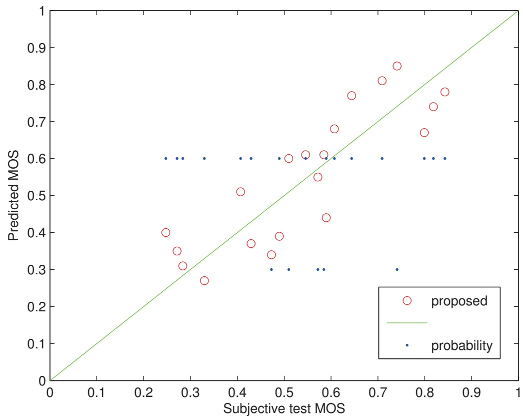

Fig. 2. Correlation between objective measurement and subjective test results.

It is quite obvious that the proposed scheme catches the experience difference among various streaming profiles while the benchmark scheme cannot.

5. CONCLUSION

We have presented in this paper a novel parametric QoE model to characterize the subjective quality of adaptive video streaming service. Our contribution lies in the innovative combination of instantaneous and cumulative quality and in the employment of an exponential decay temporal pooling model to characterize the user attention memory. The model is fundamentally different from most existing QoE assessment schemes and also features low complexity and good performance.

The proposed parametric quality model is derived based on video signal analysis and is aware of the content intensity. A subjective test is designed and performed to train and validate the proposed model. The simulation results show a good match between the output of the proposed scheme and the ground truth.

The proposed model is simple in expression and effective in performance. It can be implemented with very low complexity and serve as the utility function for a variety of streaming applications including QoE monitoring, receiver rate adaptation, and resource allocation.

Strategies outlined in Section 2 can be used as a basic guideline to build different assessment tools. In the future, more state-of-the-art VQA tools can be incorporated into the proposed QoE platform to meet specific requirements for different streaming applications.

6. REFERENCES

[1] *Information technology – Dynamic adaptive streaming over HTTP (DASH)*. ISO/IEC 23009-1:2012, 2012.

[2] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.

[3] K. Seshadrinathan and A. Bovik, “Motion tuned spatio-temporal quality assessment of natural videos,” *Image Processing, IEEE Transactions on*, vol. 19, no. 2, pp. 335–350, 2010.

[4] O. Oyman and S. Singh, “Quality of experience for HTTP adaptive streaming services,” *Communications Magazine, IEEE*, vol. 50, no. 4, pp. 20–27, 2012.

[5] K. Watanabe, J. Okamoto, and T. Kurita, “Objective video quality assessment method for evaluating effects of freeze distortion in arbitrary video scenes,” in *Image Quality and System Performance IV*, vol. Proc. SPIE 6494, Jan 2007.

[6] D. Z. Rodriguez, J. Abrahao, D. C. Begazo, R. L. Rosa, and G. Bressan, “Quality metric to assess video streaming service over TCP considering temporal location of pauses,” *Consumer Electronics, IEEE Transactions on*, vol. 58, no. 3, pp. 985–992, 2012.

[7] R. K. P. Mok, X. Luo, E. W. W. Chan, and R. K. C. Chang, “QDASH: a QoE-aware DASH system,” 2012.

[8] K. D. Singh, Y. Hadjadj-Aoul, and G. Rubino, “Quality of experience estimation for adaptive HTTP/TCP video streaming using H.264/AVC,” in *Consumer Communications and Networking Conference (CCNC)*, pp. 127–131, 2012.

[9] M. J. Kahana, *Foundation of human memory*. Oxford University Press, 2012.

[10] 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), 23rd International*, pp. 103–110, 2011.

[11] ITU, “ITU-R BT.500-13 Methodology for the subjective assessment of the quality of television pictures,” 2012.

[12] Y. Liu, Z. Li, and Y. Soh, “A novel rate control scheme for low delay video communication of H. 264/AVC standard,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 1, pp. 68–78, 2007.