Continuous Predictive Model for Quality of Experience in Wireless Video Streaming

Wenjuan Shi[1,2](✉) and Jinqiu Pan[2]

[1] College of Physics and Electronical Engineering, Yancheng Teachers University, Yancheng 224007, China
shiwj@yctu.edu.cn

[2] College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Abstract. Because of bandwidth and buffer limitation in wireless network, rebuffering events and bitrate drop often cause video impairments, e.g. compression artifacts and video stalling. Hence, these problems often make a loss of the quality of experience (QoE). For making a prediction about the impact of video impairments on QoE, a continuous predictive model for QoE in wireless video streaming is proposed. In this paper, the inputs are composed of three vectors that are the quality of video frame, rebuffering events state and human memory effect, and the output represents the predicted continuous QoE. We build the predictive model by a Hammerstein-Wiener model. Experimental results show that the proposed model can accurately make a prediction about continuous subjective QoE.

Keywords: Quality of experience (QoE) · Continuous QoE · Frame quality · Rebuffering event · Memory effect

1 Introduction

With the proliferation of smartphones, mobile video has gradually has enriched the mobile user experience. However, users are expecting to view high-quality videos by mobile smart devices. Because wireless channel has dynamic variation characteristics, it is very difficult to predict the network throughput, which can dynamically cause video transmission bitrate changing during video playing, therefore leading to affect user experience. For improving the performance of video services, it is essential to observe the quality of videos and predict the video quality for end-users.

When videos are playing, if the receiving buffer has little video data, the playing video will be interrupted to wait for new data to fill in the buffer. Such events are called rebuffering events (RE). In [1], it is shown that frequent RE can stop observers watching videos. Additionally, end users prefer video fluency to video clarity, thus ensuring video playing smoothly can efficaciously improve user experience.

For accommodating to the dynamic bandwidth changes, there are some state-of-the-art streaming techniques developed [2–7]. In the applications of wireless video streaming, quality of experience (QoE) is the finest standard to make a measurement for

G. Gui and L. Yun (Eds.): ADHIP 2019, LNICST 302, pp. 239–248, 2019.
https://doi.org/10.1007/978-3-030-36405-2_25

the quality of videos. Precise continuous predicted QoE can offer a benchmark to optimize the strategies of resource allocation in changeable wireless network.

For studying the impact factor on QoE, we consider continuous prediction for QoE as a continuous time series predictive problem, and then analyze the factors which can affect user experience, and select frame quality (FQ), rebuffering events and human memory effects (ME) to establish a predictive model, which can predict QoE accurately and supply a benchmark for evaluating the performance of video streaming control strategy.

2 Factors Impacting on QoE

It is helpful to analyze the factors impacting on QoE for understanding the influence of various events on QoE. In this section, we analyse the impact of bitrate changes, rebuffering events and memory effects on continuous QoE.

2.1 Effect of Bitrate on Continuous QoE

Because of the bandwidth dynamics in wireless network, the bitrate inevitably varies during video transmission [8]. In [9], it is shown that bitrate variation has an effect upon user perception.

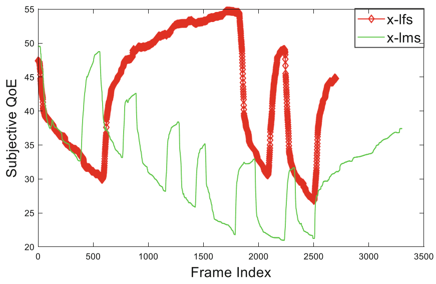
Human visual system (HVS) is highly sensitive to image edge, and has Orientation Selectivity Visual Pattern (OSVP) to extract the visual content [9, 10, 22]. In [11, 23], it is shown that each frame can be mapped into an OSVP-based histogram (OSVPH), and the extracted frame quality can be applied to make a measurement about the influence of bitrate changes on QoE. Thus we extract the quality of each frame by OSVPH, and take frame quality as one of the inputs in our proposed model.

2.2 Rebuffering Events Impacting on QoE

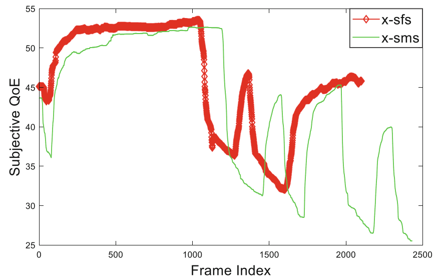
Rebuffering events (RE) can result in video stalling while observers are watching videos, which usually affects QoE [12]. For studying the impact of RE on QoE, we analyze RE impacting on QoE by the attributes of stalls, e.g. times, duration and position.

A. Stalling Times

For analyzing the impact of stalling times on QoE, we study the QoE for videos with initial delay and different times of stalls, as illustrated in Fig. 1.



(a)



(b)

Fig. 1. The QoE of a video with initial delay and different times of stalls.

Figure 1(a) represents a continuous QoE of the videos with long initial delay and few/multiple stalls (x-lfs and x-lms); Fig. 1(b) shows a continuous QoE of the videos with a short initial delay and few/multiple stalls (x-sfs and x-sms). From Fig. 1, when stalling times increases, QoE has a tendency to be smaller even when the video quality recovers to an acceptable level after stalls happen. We can observe that the stalling times can seriously affect QoE.

B. Duration of Stalls

For analyzing the impact of stalling duration on continuous QoE, supposing the initial delay and times of stalls are invariable, the subjective QoE of videos with different duration stalls is studied, as illustrated in Fig. 2. Figure 2(a) illustrates a continuous QoE of videos with short initial delay and medium stalls (x-sms) and the videos with short initial delay and long stalls (x-sls); Fig. 2(b) illustrates a continuous QoE of videos with short initial delay and short stalls (x-sss) and the videos with short initial delay and medium stalls (x-sms). It is shown that duration has a serious influence on QoE, and the stalls which last long will decrease QoE even after video quality improves.

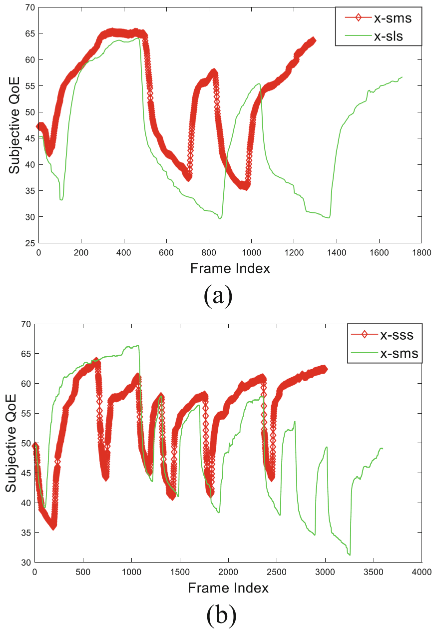


Fig. 2. The subjective QoE of videos with different stalling duration.

C. Stall Position

For better observing the impact of stall positions on continuous QoE, we use the Dynamic Time Warping (DTW) [14] to normalize the QoE of videos with different stall positions and the original reference video. The continuous QoE distribution before/after calculated by the DTW is illustrated in Fig. 3, and the DTW distance of the videos with stalls at different positions is listed in Table 1. The DTW method is one of the similarity measurement methods. While the DTW is smaller, the likelihood is greater, which means that the two time series are similar [22].

From Fig. 3 and Table 1, we can observe that no matter where the stalls occur, it will decrease the QoE, especially when continuous multi-stalls occur, the QoE will aggressively drop. In summary, we need to consider the impact of video stalls caused by RE on continuous QoE. Therefore, we take the RE state as one of the inputs in the proposed model.

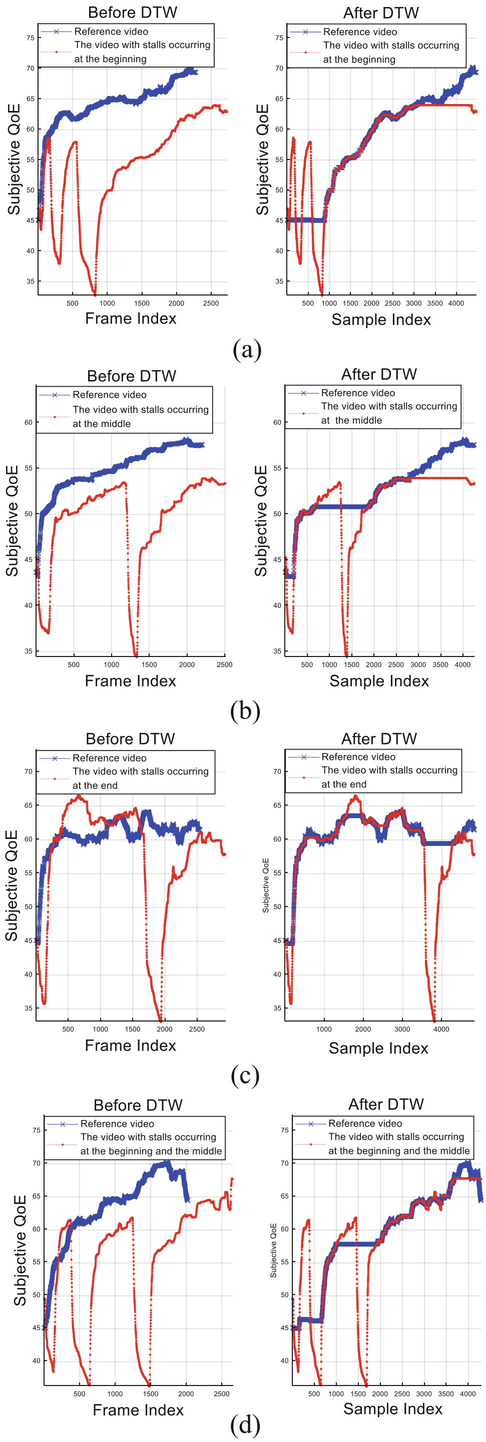
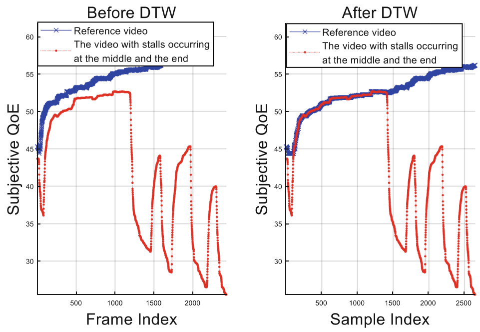
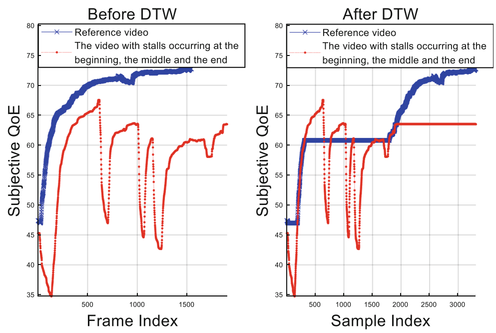


Fig. 3. The QoE of the original reference video and the videos with different stall positions before/after DTW and.



(e)



(f)

Fig. 3. (*continued*)

Table 1. The DTW of videos with stalls at different positions.

Video type	DTW distance
Videos with stalls at the beginning	3.6154
Videos with stalls at the middle	3.4256
Videos with stalls at the end	3.8518
Videos with stalls at the beginning and the middle	4.3839
Videos with continuous stalls at the middle and the end	10.0695
Videos with continuous stalls at the beginning, the middle and the end	9.0808

2.3 Human Memory Effect on Continuous QoE

When people are observing videos, they are usually affected by human memory effect (ME), e.g. recency effect, primacy effect and hysteresis effect.

(1) Recency Effect

It is well known that ME on the end part is better than others when people remember things, which is called recency effect [15, 22]. From Fig. 3 and Table 1, we can observe that the QoE of the videos with stalls at the end is a little smaller than the videos with stalls at other positions, and the DTW distance between the original reference video and the video with stalls at the end has a larger DTW value comparing with the videos with stalls at the beginning/middle. Therefore, recency effect on continuous QoE should be considered.

(2) Primacy Effect

Primacy effect is relevant to "the first impression" [16]. From Table 1, we can observe that the DTW distance between the QoE of the original reference video and that of the videos with stalls at the beginning is a little larger than the QoE of the videos with the stalls in the middle part. Thus,the stalls at the beginning has a greater impact on QoE than that at the middle. Therefore, primacy effect should also be considered.

(3) Hysteresis effect

In [17], it is shown that there is a hysteresis effect on continuous QoE during observing videos. Serious degradation of video quality during video playing impresses people negatively, and even that when video quality resumes to an acceptable level for observers, the poor impression still exists in people's memory, which causes a worse assessment for the videos.

Therefore, the impact of these memory effects on continuous QoE should not be ignored in the proposed model.

3 The Proposed Model

Considering that the Hammerstein-Wiener (HM) model can create a representation of human memory effect, we propose a block-structured nonlinear HW model. The input parameters, output parameters, and the proposed model are as follows.

A. The Inputs

From the above analysis, FQ, RE and ME are taken as the inputs of the proposed model.

(1) FQ: FQ is computed by OSVPH;
(2) RE: We define a Boolean variable RE_1, which can describe video state at time t, $RE_1 = 1$ when rebuffering events happen and $RE_1 = 0$ at the other time.
(3) ME: we take the ratio of the duration from the impairment event occurring to the end of the video to the total duration of the video as ME [22].

B. The Output

The output parameter is the predicted continuous QoE.

C. Predicted Model

The HW model is composed of one dynamic linear module and two static nonlinear modules. We describe the dynamic linear module by a transfer function with n_p poles and n_z zeros. Two static nonlinear modules represent the nonlinear relationship between the input parameters and the output parameter. The diagram of the Multiple Input Single Output (MISO) HW model is illustrated in Fig. 4.

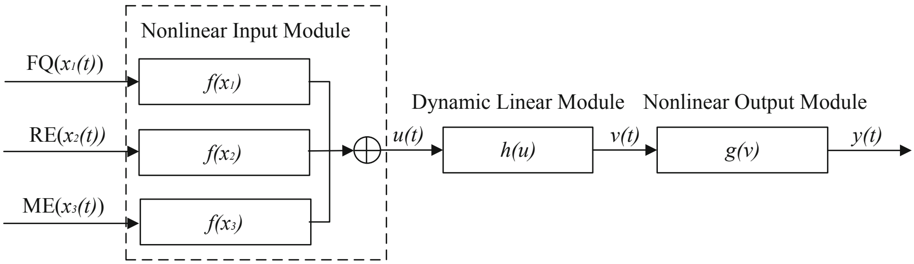


Fig. 4. The diagram of the predicted model.

In Fig. 4, z-transformation of the module $h(u)$ represents as $H(z)$. $u(t)$, $H(z)$, $y(t)$ and $f(t)$ are listed as

$$u(t) = f(x_1(t)) + f(x_2(t)) + f(x_3(t)) \tag{1}$$

$$H(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_m z^{-m}}{1 - a_1 z^{-1} - \cdots - a_n z^{-n}} \tag{2}$$

$$y(t) = g(v(t)) = \gamma_3 + \gamma_4 \frac{1}{1 + \exp(-\gamma_1 \mathrm{v}(t) + \gamma_2)} \tag{3}$$

$$f(t) = \beta_3 + \beta_4 \frac{1}{1 + \exp(-\beta_1 x_i(t) + \beta_2)} \tag{4}$$

where $x_1(t)$ is frame quality vector, $x_2(t)$ is the state vector of rebuffering events, $x_3(t)$ is the vector of memory effect, t is the frame number, $\boldsymbol{a} = [a_1, \ldots, a_n]^T$ and $\boldsymbol{b} = [b_1, \ldots, b_m]^T$ are parameter vectors and $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \gamma_3, \gamma_4]$ and $\boldsymbol{\beta} = [\beta_1, \beta_2, \beta_3, \beta_4]$ are the parameters of the sigmoid function [22].

4 The Performance Evaluation of the Proposed Model

In order to evaluate the performance of our proposed model, LIVE-Netflix mobile Video Quality Assessment database [13] is used to test the proposed model.

In this paper, we use Root-Mean-Square Error (RMSE), Outage Rate (OR) [9] and DTW [14] to assess the performance of the proposed model. RMSE can measure the difference between two time series, which can capture the overall signal fidelity. The smaller RMSE value means that the tested model has better performance. OR can measure the frequency of times when the prediction value falls outside the 95% confidence interval. The smaller OR value means the prediction accuracy of the tested model is higher [8]. DTW can quantify the similarity of two time series, the smaller DTW value means the greater similarity possibility of the two series [14]. We take PSNR, SSIM [18], MS-SSIM [19], STRRED [20], OSVPH, NIQE [21] as frame quality assessment method for FQ. The comparison results are listed in Table 2, and we mark the model with the best performance in bold text.

Table 2. Experimental comparison of median performance of the proposed model with different FR assessment methods.

FQ assessment method	RMSE	OR	DTW
PSNR	0.0825	16.8950	39.1115
SSIM	0.0739	16.3569	27.7018
MS-SSIM	0.0753	14.6119	26.1974
NIQE	0.0630	42.9224	35.1000
STRRED	0.0593	8.3019	33.9887
OSVPH	**0.0566**	**2.0111**	**27.7486**

Table 3. Performance comparison of different models.

Model	RMSE	OR	DTW
NARX	0.3322	60.9302	97.3374
GH	0.1131	47.2727	35.0410
Proposed model	**0.0566**	**2.0111**	**27.7486**

From Table 2, the performance of the proposed model taking OSVPH as FQ is obviously better than the others. Furthermore, the performance of the proposed model with the NARX model [20] and GH model [18] is compared. We list the comparison results in Table 3. We also mark the model with best performance in bold text. From Table 3, we can observe that the performance of our proposed model is obviously better than the others.

5 Conclusion

For predicting the influence of video impairment events on QoE, we proposed a continuous prediction model for Quality of Experience in wireless video streaming. The inputs are composed of FR, RE, and ME. The output is the predicted QoE. We built the proposed model by a HW MISO model. The experimental results show that the predicted result of our proposed model can make an accurate prediction about the continuous subject Quality of Experience.

References

1. Ghadiyaram, D., Pan, J., Bovik, A.C.: A subjective and objective study of stalling events in mobile streaming videos. IEEE Trans. Circuits Syst. Video Technol. **29**(1), 183–197 (2019)
2. Oyman, O., Singh, S.: Quality of experience for HTTP adaptive streaming services. IEEE Commun. Mag. **50**(4), 20–27 (2012)
3. Seufert, M., Egger, S., Slanina, M., et al.: A survey on quality of experience of HTTP adaptive streaming. IEEE Commun. Surv. Tutor. **17**(1), 469–492 (2015)
4. ISO/IEC: ISO/IEC 23009-1 Dynamics Adaptive Streaming Over HTTP (DASH) (2014)

5. Tavakoli, S., Brunnstrom, K., Gutierrez, J., et al.: Quality of experience of adaptive video streaming: investigation in service parameters and subjective quality assessment methodology. Sig. Process. Image Commun. **39**, 432–443 (2015)
6. Garcia, M.N., Simone, F.D., Tavakoli, S., et al.: Quality of experience and HTTP adaptive streaming: a review of subjective studies. In: 2014 Sixth International Workshop on Quality of Multimedia Experience, Germany, pp. 141–146 (2014)
7. Tavakoli, S., Egger, S., Seufert, M., et al.: Perceptual quality of HTTP adaptive streaming strategies: cross-experimental analysis of multi-laboratory and crowdsourced subjective studies. IEEE J. Sel. Areas Commun. **34**(8), 2141–2153 (2016)
8. Moorthy, A.K., Choi, L.K., Bovik, A.C., et al.: Video quality assessment on mobile devices: subjective, behavioral and objective studies. IEEE J. Sel. Top. Sig. Process. **6**(6), 652–671 (2012)
9. Bampis, C.G., Li, Z., Bovik, A.C.: Continuous prediction of streaming video QoE using dynamic networks. IEEE Signal Process. Lett. **24**(7), 1083–1087 (2017)
10. Wu, J., Lin, W., Shi, G.: Orientation selectivity based visual pattern for reduced-reference image quality assessment. Inf. Sci. **351**, 18–29 (2016)
11. Shi, W., Sun, Y., et al.: Spatial and temporal feature-based reduced reference quality assessment for rate-varing videos in wireless networks. Int. J. Pattern Recogn. Artif. Intell. (to appear)
12. Ghadiyaram, D., Pan, J., Bovik, A.C.: A time-varying subjective quality model for mobile streaming videos with stalling events. In: Applications of Digital Image Processing XXXVIII, pp. 1–8 (2015)
13. LIVE-Netflix Mobile Video Quality Assessment Database. http://live.ece.utexas.edu/research/LIVE_NFLXStudy/nflx_index.html
14. Berndt, D.J., Clifford, J.: Using dynamic time warping to find patterns in time series. In: AAAI 1994 Workshop on Knowledge Discovery in Databases, New York, USA, pp. 359–370 (1994)
15. Hands, D.S., Avons, S.E.: Recency and duration neglect in subjective assessment of television picture quality. Appl. Cogn. Psychol. **15**(6), 639–657 (2001)
16. Greene, J., Prepscius, C., Levy, W.B.: Primacy versus recency in quantitative model activity is the critical distinction. Learn. Mem. **7**(1), 48–57 (2000)
17. Seshadrinathan, K., Bovik, A.C.: Temporal hysteresis model of time varying subjective video quality. In: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1153–1156 (2011)
18. Wang, Z., Bovik, A.C., Sheikh, H.R., et al.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
19. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: Proceedings of Asilomar Conference on Signals, Systems, and Computers, pp. 1398–1402 (2003)
20. Soundararajan, R., Bovik, A.C.: Video quality assessment by reduced reference spatio-temporal entropic differencing. IEEE Trans. Circuits Syst. Video Technol. **23**(4), 684–694 (2013)
21. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a “completely blind” image quality analyzer. IEEE Sig. Process. Lett. **20**(3), 209–212 (2013)
22. Shi, W., Sun, Y., Pan, J.: Continuous prediction for quality of experience in wireless video streaming. IEEE Access (2019)
23. Shi, W., Sun, Y., Li, S., Cao, Q., Wang, B.: Spatial and temporal feature-based reduced reference quality assessment for rate-varying videos in wireless networks. Int. J. Pattern Recogn. Artif. Intell. (2019)