

Article

Modeling of Cumulative QoE in On-Demand Video Services: Role of Memory Effect and Degree of Interest

Tho Nguyen Duc ^{1,*}, Chanh Minh Tran ¹, Phan Xuan Tan ^{2,*}  and Eiji Kamioka ^{1,*}

¹ Graduate School of Engineering and Science, Shibaura Institute of Technology, 3 Chome-7-5 Toyosu, Koto City, Tokyo 135-8548, Japan

² SIT Research Laboratories, Shibaura Institute of Technology, 3 Chome-7-5 Toyosu, Koto City, Tokyo 135-8548, Japan

* Correspondence: mg18502@shibaura-it.ac.jp (T.N.D.); tanpx@shibaura-it.ac.jp (P.X.T.); kamioka@shibaura-it.ac.jp (E.K.)

Received: 29 June 2019; Accepted: 2 August 2019; Published: 4 August 2019



Abstract: The growing demand on video streaming services increasingly motivates the development of a reliable and accurate models for the assessment of Quality of Experience (QoE). In this duty, human-related factors which have significant influence on QoE play a crucial role. However, the complexity caused by multiple effects of those factors on human perception has introduced challenges on contemporary studies. In this paper, we inspect the impact of the human-related factors, namely perceptual factors, memory effect, and the degree of interest. Based on our investigation, a novel QoE model is proposed that effectively incorporates those factors to reflect the user's cumulative perception. Evaluation results indicate that our proposed model performed excellently in predicting cumulative QoE at any moment within a streaming session.

Keywords: quality of experience (QoE); cumulative QoE model; memory effect; degree of interest; video-on-demand services

1. Introduction

To correctly determine the end user quality of experience (QoE) for adaptive streaming services and subsequently perform a QoE based network control and management, there exists a need for the development of reliable QoE models [1] to produce a highly accurate QoE prediction either at any moment or at the end of a streaming session. This requirement links to the phenomenon of cumulative assessment where QoE can be cumulatively estimated from the time when the viewer starts watching a streaming video content to any moment of the streaming session [2]. However, such a method requires a complex quantification of multiple effects of QoE influence factors, especially human-related factors. In this paper, we propose a QoE model that precisely assesses the cumulative perceived quality in video on-demand services and it can be potentially utilized as a better alternative than either overall QoE or instantaneous QoE in QoE management.

Human-related influence factors (e.g., perceptual factors, memory effect and video content) play a crucial role in precisely modeling QoE. There have been many studies on perceived video quality models. Some studies investigate and quantify the impact of perceptual factors [3–7]. However, authors usually abandon the temporal dynamics and historical experience of the user's satisfaction, which are referred to as the memory effects [8]. Some other studies attempt to clarify the role of primacy and recency effects [9–15], resulting in the high accurate QoE prediction. Typically, the primacy and recency effects [16] determine the memory influence of impairments occurring at the beginning and the end of streaming session [17], respectively. Besides, the effect of unpleasant events which take place

in the middle of the session also leaves a considerable impact on the perceived video quality [17,18]. Theoretically, such impacts can be represented by an exponential deterioration of memory retention in time (defined by *forgetting curve*) [19–21] for infrequent events or by *repetition* [18,21] for the repeated impairments. However, the influence of forgetting behavior and repetition has not been carefully investigated in existing QoE models. Therefore, to fully express human memory effects on QoE assessment, in addition to the primacy and recency effects, the forgetting curve and repetition should be involved in the discussion. Apart from that, the factors that relate to video content also have a noticeable effect on perceived QoE. Those factors might be type of video, video complexity [22], etc. Additionally, some studies (e.g., [23–25]) have found that the user's interest in video content possibly generates the bias in his/her QoE evaluation. More concretely, the user tends to provide higher QoE scores for more attractive video contents. Such a behavior is influenced by the so-called degree of interest (DoI), which clarifies the interestingness of different video content, or the ability of the video content to attract the user and keep the user's interest [26]. However, existing studies often neglect this factor due to the fact that these numerical values might vary upon different users based on their personal interests.

In this paper, we propose a novel cumulative QoE model that extremely well quantifies multiple effects of human-related factors, that is to say, perceptual influence factors, memory effect and degree of interest (DoI). We mainly consider our model on video-on-demand streaming services since the video's DoI have to be collected offline before video streaming via performing a subjective test. To assess the accuracy in predicting cumulative QoE, our proposed model is evaluated over LFOVIA database [12] and through the subjective evaluation. Evaluation results show that the cumulative QoE at different moments within a streaming session is precisely predicted by our proposed model. Our study is distinguished with existing works by the following contributions:

- A cumulative QoE model is proposed that concurrently takes into account the human-related influence factors for predicting the time-varying cumulative QoE of on-demand streaming services.
- The novel memory weight, representing the effect of primacy, recency, forgetting and repetition is introduced in the proposed cumulative QoE model.
- The correlation between DoI and subjective QoE is investigated and confirmed in this study. Thereby, DoI becomes a potential QoE influence factor that is involved in the proposed model.

The rest of this paper is organized as follows. In Section 2, we discuss the existing works related to our approach in terms of their drawbacks. Our proposed model and our investigations on the influence factors are presented in Section 3. Section 4 evaluates the performance of the proposal and discusses the advantages and disadvantages. Section 5 concludes this paper.

2. Related Work

Modeling and predicting the user's cumulative perception to a streaming video content are able to provide lots of advantages for QoE monitor and control systems (e.g., [27,28]) since it not only reflects the user's overall satisfaction but also reveals the impact of distorted events happening during the streaming session. Most existing works only focused on modeling the overall or the instantaneous QoE, which have shown insufficient characteristics. The overall QoE [3–5,29], which demonstrates the final subjective judgment for a streaming session, can only be assessed when the viewer finishes watching. Therefore, the overall QoE cannot be applied for real-time QoE monitor and also does not give sufficient information about events occurring during the session. Although the instantaneous QoE [6,7,30], on the other hand, can provide the instant perceived video quality at a certain moment, it only reflects locally the quality assessment within a specific time range, without considering the cumulative effects of prior events. Hence, it is highly sensitive to video impairments due to hysteresis effect [18,31] and does not precisely express the user's perceived video quality. In contrast, the cumulative QoE

effectively leverages the advantages of both the overall and instantaneous QoE, while also eliminating their disadvantages.

For those reasons, the idea of cumulative assessment needs to be considered. In [2], the cumulative perceptual quality was assessed by leveraging the concept of a sliding window. The work in [32] evaluated the cumulative QoE driven by video quality, bitrate switching, and rebuffering events aligning with the exponential decay of human memory. However, these existing models did not fully express the effects of human-related influence factors (i.e., perceptual factors, primacy, recency, forgetting and repetition, and the user's interest on video content) on the video quality assessment.

For modeling QoE, there are recent studies working on cyber-physical social systems [33,34] that capture the human-related factors such as user's profiles, characteristics and interests in order to understand, predict and optimize the user's QoE. However, in the video streaming field, the impacts of human-related factors on QoE are not yet fully considered. A number of existing studies (e.g., [3–7]) have considered the perceptual factors (e.g., visual quality, rebuffering events and quality variations) without taking into account the influence of memory on subjective judgment. To support the idea of utilizing the memory effect on QoE assessment, the authors of [12,17,18] found that primacy and recency effects, which are related to the beginning and the end of a session, respectively, have significant impacts on viewer's perception. These memory effects have also been studied in [9–15], resulting in superior performances in terms of accuracy. In addition, the authors of [17,18] also indicated that those events happening in the middle of the streaming session also influence the perceived video quality. Particularly, the user tends to forget the infrequent events, but to remember the repeated one. These memory characteristics actually refer to the forgetting curve [19,21] and repetition [21]. The work in [32] considered the forgetting behavior in their cumulative QoE model, but did not employ the effect of primacy and recency. On the contrary, the authors of [2] only investigated the primacy and recency effect. Besides the above influence factors, the effect of video content was also concerned by contemporary works [9,13,23–25,35]. Accordingly, the video content (especially spatial and temporal information) plays an important role in QoE assessments. On the other hand, while the authors of [25] indicated that content type has a strong influence, the authors of [36] explored the user's satisfaction with the quality of a multimedia presentation and user's ability to analyze, synthesize and assimilate the informational content of multimedia. However, to the best of our knowledge, there is no QoE model which takes into account the user's degree of interest in video content.

In this paper, we propose a QoE model of a cumulative experience driven by human-related factors including perceptual factors, memory effect (primacy, recency and forgetting and repetition) and degree of interest.

3. Proposed Cumulative QoE Model

According to Brunnström et al. [37], QoE is defined as the results from the fulfillment of the user's expectation to the enjoyment of the application or service based on his or her *personality* and *current state*. Here, "personality" defines "the characteristics of a person that account for consistent patterns of feeling, thinking and behaving", whereas, "current state" stands for "situational or temporal changes in the feeling, thinking or behavior of a person". Therefore, the role of human-related factors in QoE modeling is extremely obvious. In this section, we first investigate the impact of those factors on human perception in QoE evaluation and then formulate our proposed cumulative QoE model.

3.1. Perceptual Factors

In video QoE assessment, perceptual factors [38] including video quality, rebuffering frequency, and rebuffering duration are directly perceived by the user. Typically, the user is usually sensitive to the current video segment quality, also known as *short time subjective quality* (STSQ) [6]. STSQ is defined as the perceptual quality of the video segment being rendered to the user. STSQ can be predicted using any of the robust video quality assessment (VQA) metrics such as Spatio-Temporal Reduced Reference Entropic Differences (STRRED) [39], Multi-Scale Structural Similarity (MS-SSIM) [40], etc.

In this study, STRRED was utilized to measure STSQ because of the exceptionally robust prediction performance [14,15]. Rebuffering occurrences also contribute a significant impact to the user’s satisfaction [18]. Therefore, rebuffering information such as rebuffering length, rebuffering position and the number of rebuffering events must be investigated. As a result, two rebuffering-related inputs employed are in this method. Firstly, *playback indicator* (PI) [11,14,15] is defined as a binary continuous-time variable, specifying the current playback status: 1 for rebuffering and 0 for normal playback. Secondly, as the user’s annoyance increases whenever a rebuffering event occurs [18], *number of rebuffering events* (NR) happening from the beginning to the current time instant of the session is considered. Intuitively, perceived video quality tends to decrease when distorted events occur, and gradually recovers since the end of those events [13]. This leads to the consideration of the fourth input which refers to the *time elapsed since the last video impairment* (TR) (i.e., bitrate switch or rebuffering occurrence) is utilized. All considered perceptual factors are fed into an LSTM-QoE model [15] to predict the instantaneous QoE as follows [15]:

$$q(t) = LSTM^0(x(t), c(t - 1)) \tag{1}$$

where $q(t)$ represents the predicted instantaneous QoE at the time instant t , $x(t)$ is the input features, and $c(t)$ is the memory cells which encode the knowledge of the inputs that have been observed up to the time t . LSTM provides two functionalities: $LSTM^0$ for output QoE prediction and $LSTM^c$ for memory cells update which is given by [15]:

$$c(t) = LSTM^c(c(0 : t - 1), q(0 : t - 1)), \forall t \geq 1 \tag{2}$$

where $c(0 : t - 1)$ and $q(0 : t - 1)$, respectively, refer to the past memory cells and the past predicted QoE.

The examples of four factors (including STSQ, PI, NR, and TR) and the architecture of LSTM-QoE model are illustrated in Figures 1 and 2, respectively.

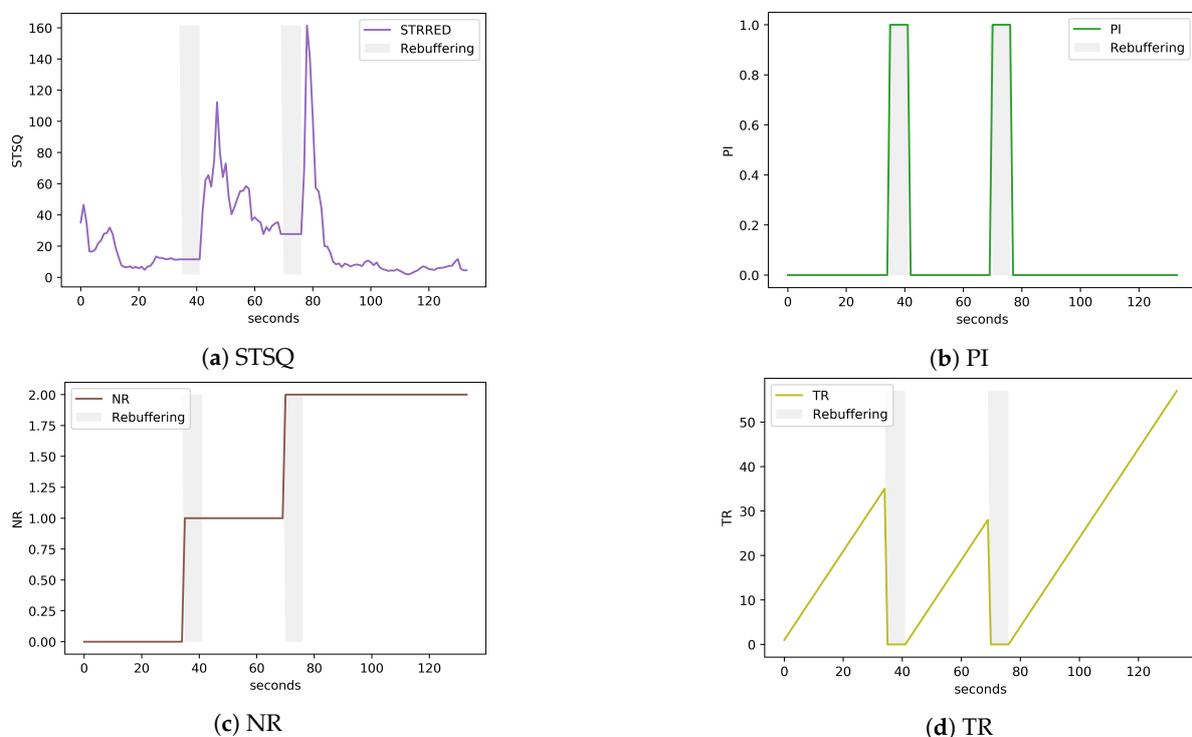


Figure 1. Example of rebuffering and bitrate-related features represented by STSQ, PI, NR, and TR.

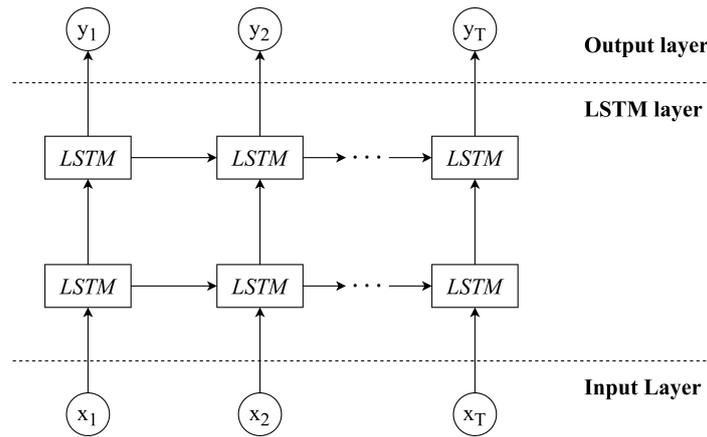


Figure 2. LSTM (Long Short-term Memory)network [15] for the user’s instantaneous QoE prediction. The network is composed of two LSTM layers. The inputs to the layers are four features including STSQ, PI, NR, and TR. The outputs combine the LSTM layers’ hidden states, representing the predicted instantaneous QoE values.

3.2. Memory Effects

Memory effects refer to the influence of historical/past experiences on the perceived video quality. Primacy and recency are two common effects which were investigated in numerous studies [11,13,15]. In addition to these factors, the effect of forgetting curve characteristic and repetition are also considered in our proposed model. The next parts of this subsection discusses the role and mathematical function of these factors. Based on that, a memory weight is proposed for the cumulative QoE model.

3.2.1. Primacy Effect

The primacy effect [16,41] describes the human behavior to recall (bitrate or rebuffering) initial events occurred at the beginning of the streaming session when providing the overall evaluation [42]. In fact, the primacy effect always exponentially decreases by time [41]. Therefore, its characteristics can be expressed by an exponential curve as follows:

$$f_P(t) = \exp(-\alpha_P * t), \quad 0 \leq t \leq L \tag{3}$$

where α_P determines the *intensity* of primacy effect (how fast the primacy effect diminishes over time) and t denotes a time instant within a session of L seconds.

3.2.2. Recency Effect

The recency effect [16,41] refers to the ability of the human memory to recall the most recent events [42]; hence, the evaluated QoE heavily depends on the recent experiences. The recency effect also can be described by an exponential curve represented by the following equation:

$$f_R(t) = \exp(-\alpha_R * (L - t)), \quad 0 \leq t \leq L \tag{4}$$

where α_R determines the *intensity* of recency effect.

The primacy effect and the recency effect can be combined as the U-shaped form [16], quantifying the influenced weight of the events occurring from the beginning to the end of a video session. As shown in Figure 3, it can be observed that Equations (3) and (4) reflect the primacy and recency effect extremely well.

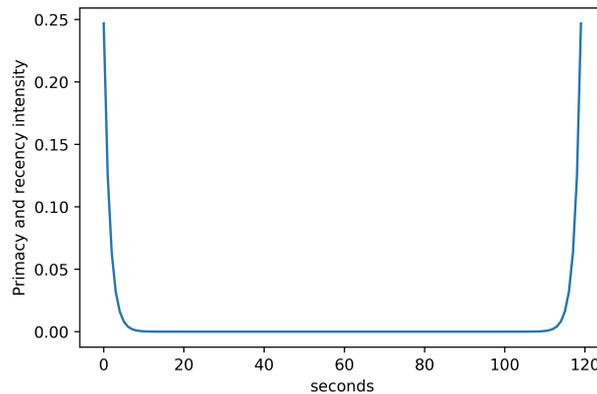
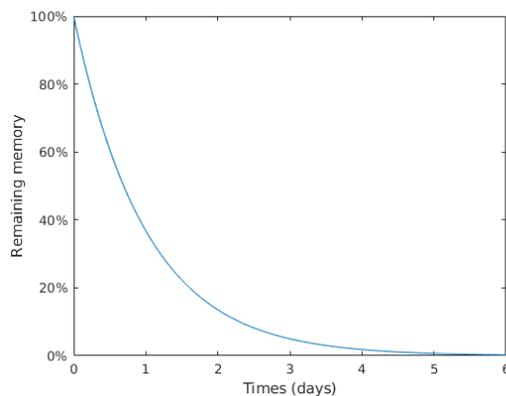


Figure 3. A typical U-shaped curve combined primacy and recency effects.

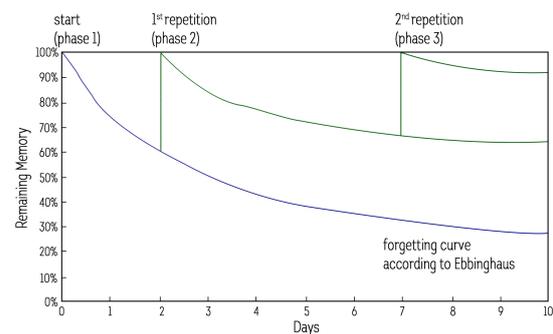
3.2.3. Forgetting Curve and Repetition

Due to the significant impact of the negative experience caused by distorted events, the primacy and recency effect can be neglected under repeated bitrate switches or rebuffering [22]. In such situations, forgetting behavior and repetition should be taken into account. The forgetting behavior, in other words, forgetting curve characteristic [21], is a natural process, describing the exponential loss of memory over time. As shown in Figure 4a, when information is learned, its memory retention declines at an exponential rate. Accordingly, any occurred events can be exponentially forgotten by time if there is no attempt to retain it. The level of remaining memory about such events at a specific time point depends on:

- The strength of memory (memory intensity): The durability that memory traces in the brain. The more annoying the event is, the stronger the user memorizes it and the longer it lasts.
- The time that has elapsed since the occurrences of events: As shown in Figure 4a, the user will forget an average of 60% of what they experience within the first period of time [20,21].
- Repetition: The more frequently an event occurs, the more likely it sticks to the user memory (as shown in Figure 4b).



(a) An example of forgetting curve



(b) Forgetting curve and repetition

Figure 4. Examples of forgetting curve and repetition.

In a typical streaming session, an interruption (bitrate switching or rebuffering) can happen regularly. When an event, especially rebuffering repeatedly occurs, the strength of memory of those events will tend to increase [21], negatively influencing the perceived video quality. Consequently, as the number of negative events increases, QoE will recover at a slower pace after the occurrence of each event. Such memory characteristics can be formulated as the following equation [43]:

$$f_{RP}(t) = \exp\left(-\frac{\alpha_{RP}}{NR(t)} * TR(t)\right), \quad 0 \leq t \leq L \tag{5}$$

where $NR(t)$ is the number of rebuffering events occurring until the time t , $TR(t)$ is the time elapsed since the last video impairment, and α_{RP} is the intensity of memory related to a rebuffering event. The ratio $\frac{\alpha_{RP}}{NR(t)}$ determines the retention of the user’s memory after the $NR(t)$ th rebuffering. Accordingly, the lower $\frac{\alpha_{RP}}{NR(t)}$ is, the higher retention rate, making f_{RP} declines at a lower rate.

3.2.4. Proposed Memory Weight

As discussed in the previous sub-subsections, the effects of primacy, recency, forgetting behavior and repetition are significantly crucial for the evaluation of the cumulative QoE. Therefore, in the proposed cumulative QoE model, we introduce a novel *memory weight* incorporating the effects of those factors to accurately assess the cumulative human perception during a streaming session. The proposed memory weight is represented by Equation (6). An example of time-varying memory weight is illustrated in Figure 5. In fact, Equation (6) is a linear combination effect of the above-mentioned memory factors obtained from Equations (3)–(5).

$$w_t = \beta_1 f_P(t) + \beta_2 f_R(t) + \beta_3 f_{RP}(t) \tag{6}$$

where $\beta_1, \beta_2, \beta_3$, respectively, determine the contribution of primacy effect, recency effect and repetition to the memory weight.

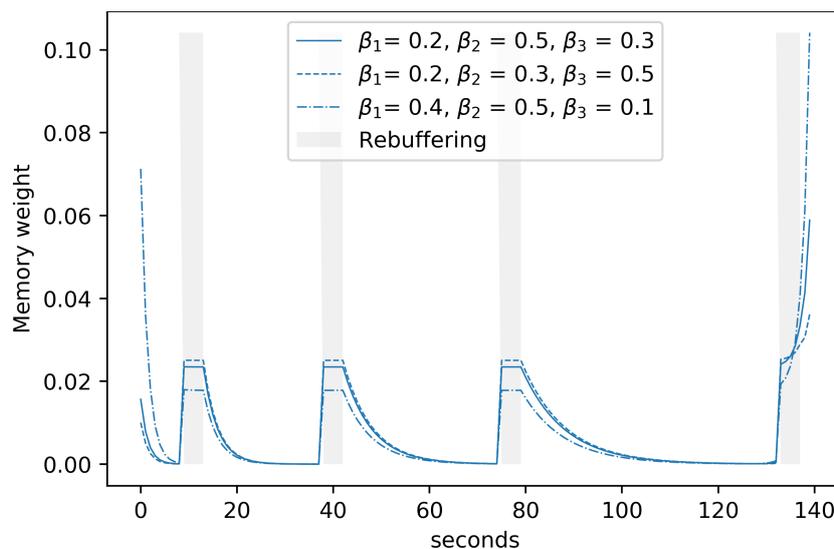


Figure 5. An example of the memory weight in a session under different values of parameters β_1, β_2 , and β_3 .

Figure 5 shows that when a rebuffering event occurs near the end of the session, the recency effect has a stronger effect on human perception. Therefore, in this period of time, the end user’s QoE will drop dramatically. In addition, the forgetting rate of a specific interruption is also smaller than those of previous ones, determining the characteristics of forgetting behavior and repetition. Therefore, the proposed memory weight potentially reflects the intensity of human memory over time during a streaming session.

3.3. Degree-of-Interest

For modeling QoE, there have been numerous studies that take into account video content-related factors (e.g., type of video, the complexity of video, etc.). However, most of them neglected the

user's interest, in other words, Degree-of-Interest (DoI). In fact, influenced by video content and viewer preferences, the user possibly has different DoI on different videos or different parts of a video. Intuitively, the user seems to provide higher QoE scores for the video with interesting content and vice versa. Typically, DoI [44] is defined as the interestingness of the video content, or the ability of the video content to attract the user and keep the user's interest [26].

To make this clear, in this study, we investigated the correlation between DoI and the overall QoE by conducting a subjective test. In this test, 18 undistorted videos from the LFOVIA Database [12] were utilized. The video content varied upon nature, wildlife, outdoor, marine, sports, animation, and gaming [12] among every video, each with a duration of 120 s. This guaranteed that the subjects would retain their interests as they watched. The referenced videos were randomly divided into six collections and encoded using FFmpeg [45] under the default settings with the resolution of 1920×1080 and were displayed on a 15-inch monitor with a resolution of 1920×1080 and a black background. The Absolute Category Rating (ACR) [46] method was used and 60 subjects agreed to participate in this experiment. Each video was assessed by at least 10 subjects. At the end of each video, the subject was asked to give an overall score representing his/her interest in the entire video content, ranging from 1 (worst or not at all interested) to 5 (best or extremely interested), following the general principle of the ITU-T recommendation P.913 [46]. A 3-min break was provided to each subject between each video to minimize the effects of viewer fatigue. The average of subjects' scores or Mean Opinion Score (MOS) for each video was utilized as the DoI of video. These values were then linearly scaled up to the range of 0–100 and compared with the corresponding overall QoE in the LFOVIA Database.

Figure 6 illustrates the obtained correlation between DoI and the overall QoE, which achieved the Pearson Correlation Coefficient (PCC) of 0.601. The correlation was modest. We speculated this was due to the small number of subjects participating in the experiment. However, it is shown that the DoI has an influence on the final decisions of the users when they provide the overall QoE. In the future, a larger number of subjects will be considered for further investigation. Based on the conclusion of this experiment, we introduce DoI as one of the potential influence factors in the proposed cumulative QoE model.

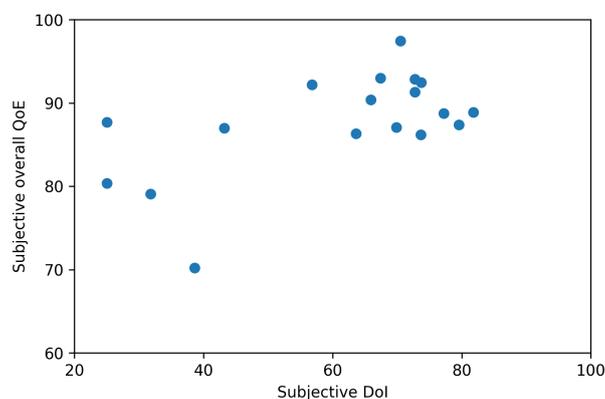


Figure 6. Scatter plot between the mean of subjective DoI scores and the subjective overall QoE obtained in the database.

3.4. Cumulative QoE Model

Through the investigation of the above human-related influence factors, the proposed cumulative QoE model is presented in Equation (7). In this model, to quantify how each of the user's past experiences influences the cumulative perception, the instantaneous QoE needs to be weighted by the memory effect from the beginning of playback to the investigated time point t within a streaming session. According to our proposed model, the procedure of estimating cumulative QoE is described as follows: Firstly, the instantaneous QoE is predicted by LSTM-QoE model [15], and stored into

vector $Q_t = (q_0, q_1, \dots, q_t)$. Secondly, the memory weight is calculated by Equation (6) to form vector $W_t = (w_0, w_1, \dots, w_t)$.

$$CQ_t = \lambda_1 (Q_t \times W_t^T) + \lambda_2 DoI \quad (7)$$

where λ_1 and λ_2 are correlation coefficients, which, respectively, determine the contribution of the user's past experience and user's interest in video content to the predicted cumulative QoE CQ_t at time instant t .

4. Performance Evaluation and Discussion

In this section, we start with the explanation of the proposed model's establishment where the necessary parameters including $\{\alpha_P, \alpha_R, \alpha_{RP}\}$, $\{\beta_1, \beta_2, \beta_3\}$ and $\{\lambda_1, \lambda_2\}$ are numerically determined. Afterward, we briefly present the evaluation and discuss the prediction performance of our model. The evaluation was two-fold. First, the prediction performance of the proposed model was quantitatively and qualitatively assessed on test videos in a specific database [12]. Second, a subjective test was conducted to evaluate how well the predicted cumulative QoE correlates with subjective cumulative evaluation at different moments of a streaming session. Finally, the complexity of the proposed model was also analyzed for real-time cumulative QoE prediction.

4.1. Model Establishment

The parameters of our proposed model were computed according to a four-step procedure as follows:

- (1) A specific publicly available database was employed for establishing and evaluating the proposed model.
- (2) An LSTM-QoE model [15] was trained to predict the instantaneous QoE values.
- (3) The memory effects' parameters $\{\alpha_P, \alpha_R, \alpha_{RP}\}$ were computed to form the memory weight vector.
- (4) The coefficients of memory weight $\{\beta_1, \beta_2, \beta_3\}$ in Equation (6) and the parameters of the proposed model $\{\lambda_1, \lambda_2\}$ in Equation (7) were determined through the predicted instantaneous QoE values and the subjective DoI collected from the experiment in Section 3.3.

The details of each step are described in the next sub-subsections.

4.1.1. Database Description

Our model was established and evaluated based on a set of 36 distorted videos in LFOVIA Video QoE Database [12]. These videos have different playout patterns distorted by bitrate switching and rebuffering events. In this database, the overall QoE and the time-varying instantaneous QoE scores for those videos were obtained are in the range $[0, 100]$, with score 0 being the worst and 100 being the best. The set of distorted videos was divided into training and testing sets with a training:testing ratio of 80:20. Accordingly, there were 28 videos in the training set and 8 videos in the testing set. The training and testing set were, respectively, used to obtain the model parameters described in Section 4.1.3 and evaluate the prediction performance of the model presented in Section 4.2.

4.1.2. Instantaneous QoE Prediction by LSTM-QoE

The instantaneous QoE values were estimated by the LSTM-QoE model [15]. The model was trained on the training set with 28 distorted videos driven by four features $STSQ$, PI , NR and TR . The performance of this model was then quantified on the eight test videos using the Pearson Correlation Coefficient (PCC) and Spearman Rank Order Correlation Coefficient (SROCC). Consequently, the model achieved high accuracy with PCC of **0.9946** and SROCC of **0.8870**. The performance of the trained model is illustrated in Figure 7, demonstrating high accurate prediction.

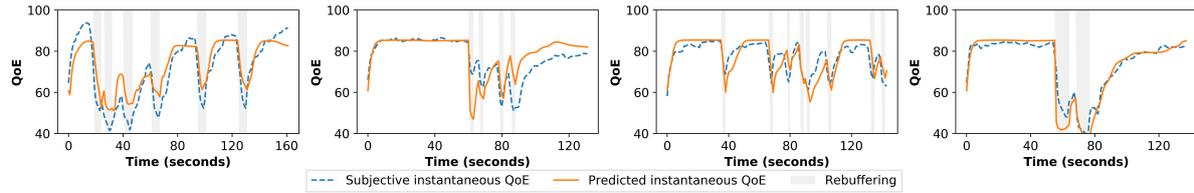


Figure 7. Some examples of instantaneous QoE prediction performance obtained from the LSTM-QoE model on different test videos of the database.

4.1.3. Parameters Selection

As discussed in Section 3.2, the parameters $\{\alpha_P, \alpha_R, \alpha_{RP}\}$ indicate how memory factors impact the perceived video quality over time. The larger $\{\alpha_P, \alpha_R, \alpha_{RP}\}$ are, the easier it is for the user to forget. According to the authors of [16,17], the effects of primacy and recency gradually decrease within 15–20 s. Therefore, in this study, the deteriorating time was set to 15 s. Since the user usually recalls unpleasant events when providing a QoE score, the effect of repetition is larger and remains longer than primacy and recency. As a result, the values of α_{RP} must be smaller than α_P and α_R . In this study, the effect of repetition remained within 30 s. The optimal remaining times of repetition effect will be analyzed in our future works. The function *solve* in MATLAB [47] was employed to compute the parameters α_P , α_R , and α_{RP} according to Equations (3)–(5), respectively. Consequently, the obtained values of parameters $\{\alpha_P, \alpha_R, \alpha_{RP}\}$ are shown in Table 1.

Table 1. Parameters of the primacy and recency effect, forgetting curve and repetition.

α_P	α_R	α_{RP}
0.6807	0.6807	0.3404

Thereby, the parameters $\{\beta_1, \beta_2, \beta_3\}$ and $\{\lambda_1, \lambda_2\}$ of weight memory and the proposed cumulative QoE model could then be estimated. Considering a streaming session with a video in the training set of L seconds, the cumulative QoE from the beginning to the end of the streaming session was calculated as follows:

$$\begin{aligned}
 CQ_L &= \lambda_1 \left(Q_L \times W_L^T \right) + \lambda_2 DoI \\
 &= \lambda_1 \sum_{i=0}^L w_i q_i + \lambda_2 DoI \\
 &= \lambda_1 \sum_{i=0}^L (\beta_1 f_P(i) + \beta_2 f_R(i) + \beta_3 f_{RP}(i)) q_i + \lambda_2 DoI
 \end{aligned} \tag{8}$$

where Q_L is the vector of instantaneous QoE (q_0, q_1, \dots, q_L) and W_L is the memory weight vector (w_0, w_1, \dots, w_L) .

As mentioned in Section 3.4, the cumulative QoE at the end of the session CQ_L was also considered as the overall QoE. Therefore, we first needed to minimize the least square error:

$$J = \|CQ_L - Q_{overall}\|^2 \tag{9}$$

where $Q_{overall}$ is the subjective overall user’s QoE obtained from the database. A curve fitting was performed using *lsqcurvefit* in MATLAB [47] with 28 training videos to obtain the memory weight parameters $\{\beta_1, \beta_2, \beta_3\}$ and the cumulative QoE parameters $\{\lambda_1, \lambda_2\}$. The numerical values of those parameters are shown in Table 2.

Table 2. Parameters of memory weight and the cumulative QoE model.

β_1	β_2	β_3	λ_1	λ_2
0.0284	0.8492	0.1177	0.9809	0.0800

4.2. Performance Evaluation on Testing Videos

After obtaining the necessary parameters for the proposed model, we quantitatively and qualitatively evaluated its prediction performance on eight distorted videos in the testing set. Alternatively, the discussion on the results was also performed.

To quantitatively assess the prediction performance, the correlation between the subjective overall QoE obtained in the LFOVIA Video QoE Database [12] and our predicted cumulative QoE at the end of each video was computed. It is crucial to note that the subjective overall QoE can be considered as the cumulative perception of the user at the end of streaming session. Three evaluation metrics were utilized for evaluation: (1) Pearson Correlation Coefficient (PCC); (2) Spearman Rank Order Correlation Coefficient (SROCC); and (3) Root Mean Square Error (RMSE). Typically, PCC and SROCC quantify how well the predicted QoE tracks the actual QoE scores in the database, whereas RMSE indicates the closeness between them. We also compared our proposed model with a reference method in [32], using the same training set and testing set. The cumulative QoE model in [32] is characterized by the following equation:

$$Q_t = \gamma Q_{t-1} + (1 - \gamma)q_t \quad (10)$$

where q_t is the instantaneous user experience at moment t , Q_{t-1} is the cumulative QoE at the previous moment $t - 1$, and γ is the memory strength parameter. The correlation between the predicted cumulative QoE obtained from this model and the subjective overall QoE in LFOVIA database was also investigated through PCC, SROCC and RMSE metrics. We report the performance of our model and the reference method in Table 3. This result shows a superior prediction performance of our model. Figure 8 additionally emphasizes the competitive performance of our model. Thereby, the proposed model has effectively assessed cumulative perception over multiple scenarios in testing videos.

In qualitative evaluation, our purpose was to validate the impact of memory effects and DoI on the cumulative QoE prediction over multiple scenarios on testing videos. Thereby, the prediction performance of the proposed model in a short period and longer period could be assessed. Thus, in Figure 9, we plot the predicted cumulative QoE in comparison with both subjective instantaneous QoE and subjective overall QoE that were obtained from the database. Hereafter, the terms of subjective instantaneous QoE and subjective overall QoE are referred to as instantaneous QoE and overall QoE, respectively.

In general, the predicted cumulative QoE precisely reacts to any interruption at any moment while being close to the overall QoE at the end of the streaming session. For the initial interruption, it is always witnessed a significant deterioration in predicted cumulative QoE. Nevertheless, when such unpleasant events continuously occur, the predicted cumulative QoE tends to decrease at a lower rate. Additionally, a lower recovering rate is subsequently introduced after each event.

Table 3. Prediction performance of the reference model and our proposed model over training and testing set.

		PCC	SROCC	RMSE
Training	[32]	0.7413	0.6420	10.6187
	Proposed model	0.9441	0.8604	4.1525
Testing	[32]	0.2777	0.2381	7.5135
	Proposed model	0.7664	0.7857	4.6538

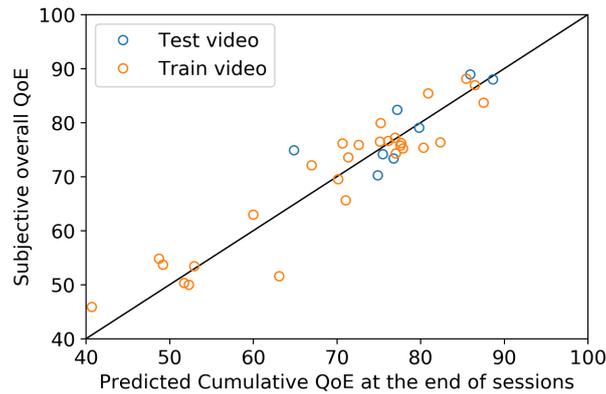


Figure 8. Correlation between subjective overall QoE and predicted cumulative QoE at the end of streaming session.

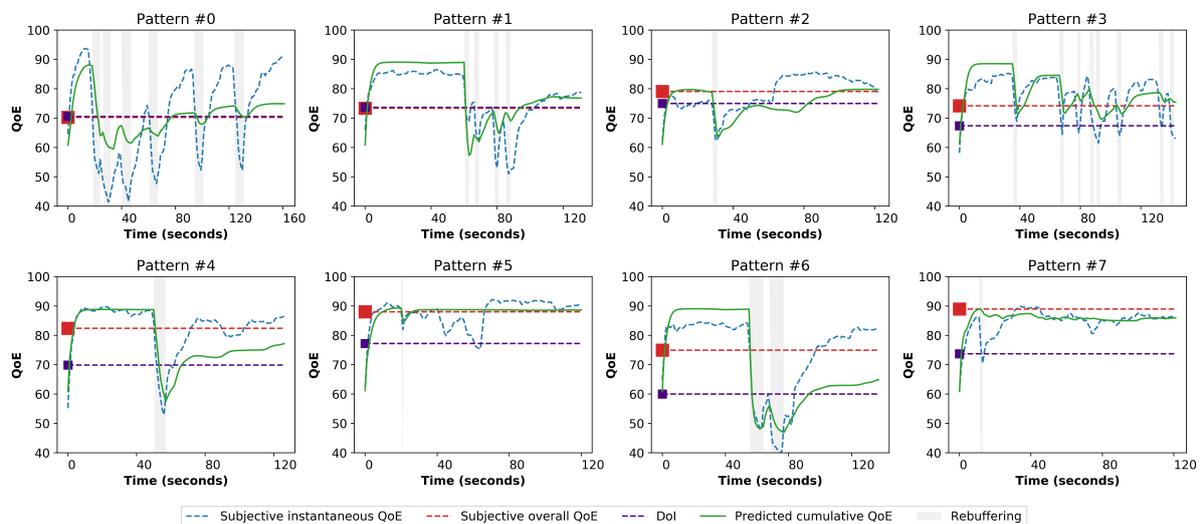


Figure 9. Predicted cumulative QoE in comparison with the subjective overall and instantaneous QoE over eight different playout patterns.

4.2.1. Impacts of Memory Effects

In Patterns #2, #5, and #7, there is only one interruption with short duration occurring near the beginning of streaming sessions. As a result, the predicted cumulative QoE introduces a slight decrease, followed by a gradual recovery and convergence to a value as close to the overall QoE. These trends are consistent with those of instantaneous QoE. It means that the prediction accurately demonstrates the role of forgetting curve characteristic as well as the recency effect. More concretely, after the interruption finishes, the memory intensity about such event starts to exponentially decay, leading to the recovery in perceived video quality. At the end of streaming sessions, there is a possibility that the decay has completely finished; in other words, the memory of distorted events has vanished. Therefore, the recency effect becomes dominant, leading to the consistency among the predicted cumulative QoE, instantaneous QoE, and overall QoE.

In Patterns #0 and #1, rebuffering event repeatedly occurs in the middle of streaming sessions. While the predicted cumulative QoE is consistent with the overall QoE at the end of sessions, the instantaneous QoE tends to continuously increase, creating a big gap to the overall QoE. At first sight, one might think that the overall QoE must be as high as the instantaneous QoE at the end of streaming sessions due to the recency effect. This inference is understandable because the moment at which the last interruption occurs is quite far from the end of sessions, thus the recency effect would have become dominant, resulting in the consistency among those QoE evaluations. However,

when the interruption repeats many times, the impact of repetition characteristic become significantly obvious. Consequently, the user tends to provide an overall evaluation whose value is lower than the instantaneous QoE. On the other hand, by considering the recency effect and repetition characteristic, our proposed model can effectively provide the prediction consistency with the overall QoE.

According to the hysteresis effect [31], the user is highly sensitive to a single unpleasant event and provides poor QoE scores immediately. However, when the interruption occurs many times as in Patterns #0, #1 and #3, the impact of the hysteresis effect will be shared with the repetition characteristic. This makes the user behave in the consideration of past annoying events to avoid the aggressive reaction. In addition, under the impact of repetition characteristic, such events are stuck in the user's memory and are recalled when the user provides the overall assessment. However, the instantaneous QoE always aggressively reacts to the distorted events, by dramatically decreasing and quickly recovering during a short period. This is because the instantaneous QoE is estimated locally without considering the global views of the streaming session. Oppositely, by weighting the instantaneous QoE by the memory effects (especially repetition characteristic), the predicted cumulative QoE can react calmly, and, eventually, correlates perfectly with the overall QoE.

Interestingly, the predicted cumulative QoE also indicates a special behavior in human perception, which cannot be found in the instantaneous QoE and overall QoE. We call such the behavior as the *persistent evaluation* where the user seems familiar with the distorted event and accepts it. The user does not even want to deteriorate their evaluation score or to quit from the streaming session. For instance, Patterns #0 and #3 visualize that the cumulative QoE dramatically falls after the occurrence of the first rebuffering event. However, it decreases with a significantly lower amplitude on the ones happening subsequently.

4.2.2. Impacts of DoI

As mentioned in Section 3.3, the correlation between DoI and subjective overall QoE is modest. However, the contribution of DoI on prediction performance is well recognized in some cases, as shown in Patterns #2, #3, #5 and #7, which share a common characteristic where the predicted cumulative QoE correctly meets the overall QoE. Without DoI ($\lambda_2 = 0$), the predicted cumulative QoE would have been much lower than the overall QoE. Especially, in Patterns #5 and #7, which contain only one super-short rebuffering event near the beginning of streaming sessions, the memory intensity about this event must have completely vanished, followed by the dominance of the recency effect, resulting in very high overall QoE. However, the contents of these two videos might not be sufficiently interesting to the users, leading to the deterioration in their evaluation. Therefore, when the contribution of DoI is precisely recognized, our proposed model provides an extremely high accurate prediction. However, in Patterns #4 and #6, there exist long duration interruptions in the middle of streaming sessions, creating significantly high intensity memory about those events. As a result, the predicted cumulative QoE dramatically decreases and slowly recovers. However, the insufficiently accurate contribution of DoI has curbed the recovering rate. Consequently, the predicted cumulative QoE cannot catch up with the overall QoE at the end of streaming sessions. This emphasizes the lack of generalization in DoI coefficient λ_2 . We believe that the original reason is the insufficient number of participated subjects in the subjective evaluation in Section 3.3 where each video was watched and evaluated by only 10 subjects. In the future, a larger number of participants must be involved in this experiment.

4.3. Subjective Evaluation

A subjective evaluation was conducted to assess the accuracy of the proposed model aligning with ground truth QoE scores provided by a number of subjects. The performance of QoE prediction using the proposed model was evaluated by relying on the following four measures: (1) PCC; (2) SROCC; (3) RMSE; and (4) Outage Rate (OR) [6]. While PCC and SROCC quantify the correlation between predicted cumulative QoE and the subjective cumulative QoE, the closeness between predicted scores and the ground truth scores is numerically obtained by using RMSE and OR. In particular, OR which

measures the frequency of times when the prediction p_i falls outside twice the confidence interval of subjective scores s_i , is defined as the following equation:

$$OR = \frac{1}{N} \sum_i^N \mathbb{1}(|p_i - s_i| > 2CI_{s_i}) \quad (11)$$

where $\mathbb{1}(\cdot)$ is the indicator function

To conduct the subjective test, six distorted videos from the testing set of LFOVIA database (Patterns #0, #1, #3, #4, #5, and #7) were selected. We selected those videos because they have different contents, thus the role of DoI in our model could potentially be assessed. Each distorted video was cropped into four small videos with starting timestamps of 00:00:00 and different length (60, 80, 100, and 120 s) using FFmpeg [45]. The purpose was to ask the subjects to provide subjective cumulative evaluations at the time points of 60, 80, 100, and 120 s of each distorted video. The correlations between subjective cumulative QoE and predicted cumulative QoE obtained from our model and reference model were assessed. The cropped videos were divided into six collections with different video content and displayed on a 15-inch screen with a resolution of 1920×1080 and a black background. Each video was rated by at least 18 subjects and there were totally 120 participants. Note that these subjects were different from those in the “DoI” experiment presented in Section 3.3. The Absolute Category Rating method was used in our experiment [46]. The subjects give a rating score at the end of each cropped video with the score ranging from 1 (worst) to 5 (best) based on the perceived quality and video content, following the general principle of the ITU-T recommendation P.913 [46]. The average of subjects’ scores, associated with 95% confidence interval, for each cropped video, was utilized as the subjective cumulative QoE. These values were linearly rescaled so that the scores lay in the range 0–100 and then compared with the predicted cumulative QoE.

Figure 10 illustrates the obtained correlation between the predicted cumulative QoE and subjective cumulative QoE. The comparison in QoE prediction performance between our model and reference model is tabulated in Table 4. Accordingly, we observed that the proposed model provides a competitive performance in terms of SROCC, RMSE and OR against the reference model. On the other hand, Figure 11 shows a reasonable prediction performance of our model in comparison with subjective cumulative QoE at four discrete moments (at time points of 60, 80, 100, and 120 s) within a streaming session. In general, the proposed model performs extremely well when the high frequent and long duration rebuffering occur. It means that our model is capable of cumulatively capturing the effects of all the occurred unpleasant events on human perception. However, the model performances on Patterns #5 and #7 are poorer, as compared to other patterns (Patterns #0, #1, and #4) even though they have only one short rebuffering event. This is because, on Patterns #5 and #7, the users’ perception seems to be significantly affected by the video content. In other words, the effect of DoI becomes dominant in their evaluation, which is not precisely captured by our model.

Table 4. Prediction performance of reference model and the proposed model over subjective experiment.

	PCC	SROCC	RMSE	OR (%)
[32]	0.5418	0.3917	9.1318	33.3
Proposed model	0.5405	0.5146	9.0922	25.0

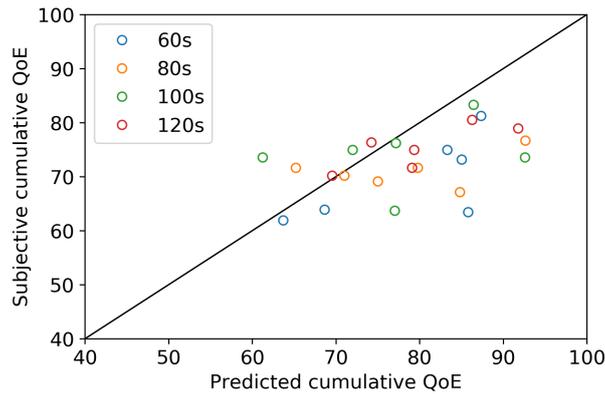


Figure 10. Scatter plot of predicted cumulative QoE and subjective cumulative QoE.

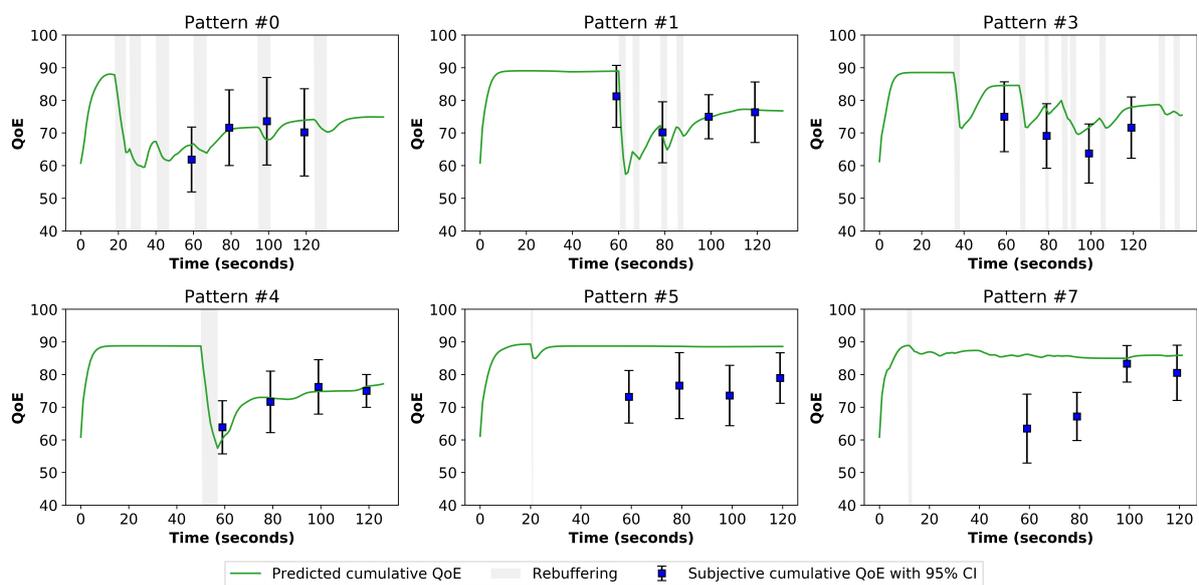


Figure 11. Performance of our predicted cumulative QoE in comparison with the subjective cumulative QoE.

4.4. Computational Complexity

The computational complexity of our proposed model was determined by the computational complexity of forming the instantaneous QoE vector $Q_t = (q_0, q_1, \dots, q_t)$ predicted by the LSTM-QoE model. It is important to note that, at the time instant t , the previous instantaneous QoE values $\{q_0, q_1, \dots, q_{t-1}\}$ have already been predicted and cached in the memory. Since the LSTM-QoE model takes up only a very small computational overhead to predict q_t , to form the vector Q_t , the cumulative QoE of each second CQ_t can be predicted in real-time. To demonstrate this, we calculated the required computing time for training LSTM-QoE model and predicting the instantaneous QoE at the end of a session q_L . All the timing experiments were carried out on a 18.04 Ubuntu LTS Intel i7-8750H @ 2.20 GHz and 16 GB RAM system. The LSTM-QoE model took 620.740 s to train and 0.4917 ms to predict q_L . Furthermore, the cumulative QoE CQ_L prediction took 0.5103 ms. Thus, our proposed model is suitable for real-time cumulative QoE prediction.

4.5. Overall Evaluation

We assessed the performance of our proposed model on a publicly available database and the subjective test. In this way, we could validate the predicted cumulative QoE in both quantitative and qualitative manners. Typically, the model can precisely provide cumulative QoE prediction in

different scenarios. Therefore, our proposal promisingly provides an alternative and reliable approach in modeling QoE towards QoE based control and management.

5. Conclusions and Future Work

In this paper, a novel cumulative QoE model is proposed. This model successfully and effectively incorporates the impacts of human-related influence factors to predict the cumulative perceived video quality. In different scenarios, the proposed model achieved impressive performance, outperforming the reference model. Additionally, it was shown that the introduced memory weight accurately mimicked human memory during a streaming session, especially when unpleasant events repeatedly occurred. Besides, the user's interest in video content was found as a potential influence factor in predicting QoE. However, the correlation between DoI and subjective overall QoE was not so high due to the small number of subjects involved in our experiment. For future work, the influence of DoI will be further investigated. In addition, the proposed cumulative QoE model will also be evaluated in multiple databases to understand how well the model will perform across diverse scenarios of video streaming.

Author Contributions: Conceptualization, T.N.D., C.M.T. and P.X.T.; methodology, T.N.D., C.M.T. and P.X.T.; writing—original draft preparation, T.N.D. and C.M.T.; writing—review and editing, P.X.T. and E.K.; and supervision, E.K. and P.X.T.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Barman, N.; Martini, M.G. QoE Modeling for HTTP Adaptive Video Streaming—A Survey and Open Challenges. *IEEE Access* **2019**, *7*, 30831–30859. [[CrossRef](#)]
2. Tran, H.T.T.; Ngoc, N.P.; Hoßfeld, T.; Thang, T.C. A Cumulative Quality Model for HTTP Adaptive Streaming. In Proceedings of the 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), Sardinia, Italy, 29 May–1 June 2018; pp. 1–6. [[CrossRef](#)]
3. De Vriendt, J.; De Vleeschouwer, D.; Robinson, D. Model for estimating QoE of video delivered using HTTP adaptive streaming. In Proceedings of the 2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013), Ghent, Belgium, 27–31 May 2013; pp. 1288–1293.
4. Liu, Y.; Dey, S.; Ulupinar, F.; Luby, M.; Mao, Y. Deriving and Validating User Experience Model for DASH Video Streaming. *IEEE Trans. Broadcast.* **2015**, *61*, 651–665. [[CrossRef](#)]
5. Zegarra Rodríguez, D.; Lopes Rosa, R.; Costa Alfaia, E.; Issy Abrahão, J.; Bressan, G. Video Quality Metric for Streaming Service Using DASH Standard. *IEEE Trans. Broadcast.* **2016**, *62*, 628–639. [[CrossRef](#)]
6. Chen, C.; Choi, L.K.; de Veciana, G.; Caramanis, C.; Heath, R.W.; Bovik, A.C. Modeling the Time-Varying Subjective Quality of HTTP Video Streams With Rate Adaptations. *IEEE Trans. Image Process.* **2014**, *23*, 2206–2221. [[CrossRef](#)] [[PubMed](#)]
7. Garcia, M.N.; Robitza, W.; Raake, A. On the accuracy of short-term quality models for long-term quality prediction. In Proceedings of the 2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX), Costa Navarino, Messinia, Greece, 26–29 May 2015; pp. 1–6. [[CrossRef](#)]
8. Hoßfeld, T.; Biedermann, S.; Schatz, R.; Platzer, A.; Egger, S.; Fiedler, M. The memory effect and its implications on Web QoE modeling. In Proceedings of the 2011 23rd International Teletraffic Congress (ITC), San Francisco, CA, USA, 6–9 September 2011.
9. Shen, Y.; Liu, Y.; Liu, Q.; Yang, D. A method of QoE evaluation for adaptive streaming based on bitrate distribution. In Proceedings of the 2014 IEEE International Conference on Communications Workshops (ICC), Sydney, Australia, 10–14 June 2014; pp. 551–556. [[CrossRef](#)]
10. Bampis, C.G.; Bovik, A.C. Learning to predict streaming video QoE: Distortions, rebuffering and memory. *arXiv* **2017**, arXiv:1703.00633.
11. Bampis, C.G.; Li, Z.; Bovik, A.C. Continuous Prediction of Streaming Video QoE Using Dynamic Networks. *IEEE Signal Process. Lett.* **2017**, *24*, 1083–1087. [[CrossRef](#)]

12. Eswara, N.; Manasa, K.; Kommineni, A.; Chakraborty, S.; Sethuram, H.P.; Kuchi, K.; Kumar, A.; Channappayya, S.S. A Continuous QoE Evaluation Framework for Video Streaming Over HTTP. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *28*, 3236–3250. [[CrossRef](#)]
13. Ghadiyaram, D.; Pan, J.; Bovik, A.C. Learning a Continuous-Time Streaming Video QoE Model. *IEEE Trans. Image Process.* **2018**, *27*, 2257–2271. [[CrossRef](#)]
14. Eswara, N.; Sethuram, H.P.; Chakraborty, S.; Kuchi, K.; Kumar, A.; Channappayya, S.S. Modeling Continuous Video QoE Evolution: A State Space Approach. In Proceedings of the 2018 IEEE International Conference on Multimedia and Expo (ICME), San Diego, CA, USA, 23–27 July 2018; pp. 1–6. [[CrossRef](#)]
15. Eswara, N.; Ashique, S.; Panchbhai, A.; Chakraborty, S.; Sethuram, H.P.; Kuchi, K.; Kumar, A.; Channappayya, S.S. Streaming Video QoE Modeling and Prediction: A Long Short-Term Memory Approach. *IEEE Trans. Circuits Syst. Video Technol.* **2019**. [[CrossRef](#)]
16. Greene, A.J. Primacy Versus Recency in a Quantitative Model: Activity Is the Critical Distinction. *Learn. Mem.* **2000**, *7*, 48–57. [[CrossRef](#)]
17. Bampis, C.G.; Li, Z.; Moorthy, A.K.; Katsavounidis, I.; Aaron, A.; Bovik, A.C. Study of Temporal Effects on Subjective Video Quality of Experience. *IEEE Trans. Image Process.* **2017**, *26*, 5217–5231. [[CrossRef](#)]
18. 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.* **2019**, *29*, 183–197. [[CrossRef](#)]
19. Bampis, C.G.; Li, Z.; Katsavounidis, I.; Huang, T.Y.; Ekanadham, C.; Bovik, A.C. Towards Perceptually Optimized End-to-end Adaptive Video Streaming. *arXiv* **2018**, arXiv:1808.03898.
20. Loftus, G.R. Evaluating forgetting curves. *J. Exper. Psych. Learn. Mem. Cognit.* **1985**, *11*, 397–406. [[CrossRef](#)]
21. Ebbinghaus, H. Memory: A contribution to experimental psychology. *Ann. Neurosci.* **2013**, *20*, 155–156. [[CrossRef](#)]
22. Hoßfeld, T.; Seufert, M.; Sieber, C.; Zinner, T. Assessing effect sizes of influence factors towards a QoE model for HTTP adaptive streaming. In Proceedings of the 6th International Workshop on Quality of Multimedia Experience, QoMEX 2014, Singapore, 18–20 September 2014; pp. 111–116. [[CrossRef](#)]
23. Ghinea, G.; Thomas, J.P. QoS Impact on User Perception and Understanding of Multimedia Video Clips. In Proceedings of the Sixth ACM International Conference on Multimedia, Bristol, UK, 13–16 September 1998; ACM: New York, NY, USA, 1998; pp. 49–54. [[CrossRef](#)]
24. Lee, J.S.; De Simone, F.; Ebrahimi, T. Subjective quality evaluation VIA paired comparison: Application to scalable video coding. *IEEE Trans. Multimed.* **2011**, *13*, 882–893. [[CrossRef](#)]
25. Ries, M.; Froehlich, P.; Schatz, R. QoE evaluation of high-definition IPTV services. In Proceedings of the 21st International Conference, Radioelektronika Brno, Czech Republic, 19–20 April 2011; pp. 15–20. [[CrossRef](#)]
26. Le Callet, P.; Benois-Pineau, J. Visual Content Indexing and Retrieval with Psycho-Visual Models. In *Visual Content Indexing and Retrieval with Psycho-Visual Models, Multimedia Systems and Applications*; Springer: Cham, Switzerland, 2017; pp. 1–10.
27. Kassler, A.; Skorin-Kapov, L.; Dobrijevic, O.; Matijasevic, M.; Dely, P. Towards QoE-driven multimedia service negotiation and path optimization with software defined networking. In Proceedings of the 20th International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2012), Split, Croatia, 11–13 September 2012; pp. 1–5.
28. Ben Letaifa, A. Adaptive QoE monitoring architecture in SDN networks: Video streaming services case. In Proceedings of the 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia, Spain, 26–30 June 2017; pp. 1383–1388. [[CrossRef](#)]
29. Orsolich, I.; Pevec, D.; Suznjevic, M.; Skorin-Kapov, L. A machine learning approach to classifying YouTube QoE based on encrypted network traffic. *Multimed. Tools Appl.* **2017**, *76*, 22267–22301. [[CrossRef](#)]
30. Duc, T.N.; Tran, C.M.; Tan, P.X.; Kamioka, E. Bidirectional LSTM for Continuously Predicting QoE in HTTP Adaptive Streaming. In Proceedings of the 2019 2nd International Conference on Information Science and Systems (ICISS 2019), Tokyo, Japan, 16–19 March 2019; ACM: New York, NY, USA, 2019; pp. 156–160. [[CrossRef](#)]
31. Seshadrinathan, K.; Bovik, A.C. Temporal hysteresis model of time varying subjective video quality. In Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011; pp. 1153–1156. [[CrossRef](#)]

32. Xue, J.; Zhang, D.-Q.; Yu, H.; Chen, C.W. Assessing quality of experience for adaptive HTTP video streaming. In Proceedings of the 2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), Chengdu, China, 14–18 July 2014; pp. 1–6. [[CrossRef](#)]
33. Tsiropoulou, E.E.; Thanou, A.; Papavassiliou, S. Modelling museum visitors' Quality of Experience. In Proceedings of the 2016 11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), Thessaloniki, Greece, 20–21 October 2016; pp. 77–82. [[CrossRef](#)]
34. Tsiropoulou, E.E.; Thanou, A.; Papavassiliou, S. Quality of Experience-based museum touring: A human in the loop approach. *Soc. Netw. Anal. Mining* **2017**, *7*, 1–13. [[CrossRef](#)]
35. Duanmu, Z.; Ma, K.; Wang, Z. Quality-of-Experience for Adaptive Streaming Videos: An Expectation Confirmation Theory Motivated Approach. *IEEE Trans. Image Process.* **2018**, *27*, 6135–6146. [[CrossRef](#)]
36. Gulliver, S.R.; Ghinea, G. Stars in their eyes: What eye-tracking reveals about multimedia perceptual quality. *IEEE Trans. Syst. Man Cybernet. Part A Syst. Hum.* **2004**, *34*, 472–482. [[CrossRef](#)]
37. Brunnström, K.; Beker, S.A.; De Moor, K.; Dooms, A.; Egger, S.; Garcia, M.N.; Hossfeld, T.; Jumisko-Pyykkö, S.; Keimel, C.; Larabi, M.C.; et al. Qualinet white paper on definitions of quality of experience. *Archive ouverte HAL* **2013**, hal-00977812.
38. Seufert, M.; Egger, S.; Slanina, M.; Zinner, T.; Hoßfeld, T.; Tran-Gia, P. A Survey on Quality of Experience of HTTP Adaptive Streaming. *IEEE Commun. Surv. Tutor.* **2015**, *17*, 469–492. [[CrossRef](#)]
39. Soundararajan, R.; Bovik, A.C. Video Quality Assessment by Reduced Reference Spatio-Temporal Entropic Differencing. *IEEE Trans. Circuits Syst. Video Tech.* **2013**, *23*, 684–694. [[CrossRef](#)]
40. Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multiscale structural similarity for image quality assessment. In Proceedings of the Thrity-Seventh Asilomar Conference on Signals, Systems Computers, Pacific Grove, CA, USA, 9–12 November 2003; Volume 2, pp. 1398–1402. [[CrossRef](#)]
41. Murdock, B.B., Jr. The serial position effect of free recall. *J. Exp. Psychol.* **1962**, *64*, 482–488. [[CrossRef](#)]
42. Rundus, D. Analysis of rehearsal processes in free recall. *J. Exp. Psychol.* **1971**, *89*, 63–77. [[CrossRef](#)]
43. Murakowski, J.; Wozniak, P.; Gorzelanczyk, E. Two components of long-term memory. *Acta Neurobiol. Exp.* **1995**, *55*, 301–305.
44. Hoßfeld, T.; Schatz, R.; Egger, S. SOS: The MOS is not enough! In Proceedings of the 2011 3rd International Workshop on Quality of Multimedia Experience (QoMEX 2011), Mechelen, Belgium, 7–9 September 2011; pp. 131–136. [[CrossRef](#)]
45. Bellard, F. FFmpeg Multimedia System. Available online: <https://www.ffmpeg.org/about.html> (accessed on 22 June 2019).
46. ITU-T. Methods for the Subjective Assessment of Video Quality, Audio Quality and Audiovisual Quality of Internet Video and Distribution Quality Television in any Environment. *Recommendation ITU-T P.913* **2016**. [[CrossRef](#)]
47. *MATLAB, version 9.6.0 (R2019a)*; The MathWorks Inc.: Natick, MA, USA, 2019.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).