Master's Thesis

Title

Real-time QoE estimation method using EEG for video delivery services

Supervisor

Professor Masayuki Murata

Author

Kasumi Kitao

January 29th, 2021

Department of Information Networking

Graduate School of Information Science and Technology

Osaka University

Master's Thesis

Real-time QoE estimation method using EEG for video delivery services

Kasumi Kitao

Abstract

Video streaming services such as YouTube, Netflix, and Hulu, and remote web conferencing systems such as Zoom and Cisco WebEx are rapidly gaining popularity, and the number of users of such services continues to increase every year. Research on improving the quality of user experience (QoE) in video streaming delivery (hereafter referred to as video delivery), which is common to these services, has become very important for both users and service providers. In recent years, most video delivery services have adopted HTTP adaptive streaming (HAS) in order to provide video delivery with appropriate service quality for various service usage environments. Dynamic Adaptive Streaming over HTTP (MPEG-DASH) is one of the most popular HAS standards, and there has been a lot of research on developing QoE-based adaptive streaming using MPEG-DASH. QoE is expected to be used as an indicator to reflect user satisfaction. However, QoE depends on various user's external factors and internal factors. In our previous research, we made a classifier for estimating QoE by using extracting temporal and frequency features from the user's EEG, and clarified that the estimation accuracy of the method. In the context of the QoE-based video delivery control, it is important to accurately estimate the user's QoE especially when the video quality changes. Therefore, in this thesis, in addition to our previous method, we investigate a method to obtain the timing for controlling the video bitrate. We propose a delivery control method based on the user QoE, using event-related potentials (ERPs), which are EEGs occurred by a specific event (stimulus) and are endogenous potentials that reflect the subject's cognitive attitude toward the stimulus. The P300 is an event-related potential and is known to occur in association with the presentation of low-frequency stimuli in stimulus-classification tasks such as the oddball task, but recent studies have reported that it is also evoked when perceiving changes in a playback video quality. We collected the EEG and QoE data during video watching, and evaluated the value of recall when bad label is classified as target, and showed that the recall value was 76.0% at maximum and the average is 49.3%. And We evaluated the P300 on a threshold basis, detecting and comparing it to the bitrate changes. We also discussed the design and feasibility of a rate control system using an EEG-based QoE estimation method.

Keywords

Quality of Experience (QoE), Event Related Potentials (ERP),
Adaptive Streaming, Genetic Algorithm (GA)

Contents

List of Figures

List of Tables

1 Introduction

Video streaming services such as YouTube, Netflix, and Hulu, and remote web conferencing systems such as Zoom and Cisco WebEx are rapidly gaining popularity, and the number of users of such services continues to increase every year. Research on improving the quality of user experience (QoE) in video streaming delivery (hereafter referred to as video delivery), which is common to these services, has become very important for both users and service providers. In recent years, most video delivery services have adopted HTTP adaptive streaming (HAS) in order to provide video delivery with appropriate service quality for various service usage environments. Dynamic Adaptive Streaming over HTTP (MPEG-DASH) is one of the most popular HAS standards, and there has been a lot of research on developing QoE-based adaptive streaming using MPEG-DASH [1, 2]. These studies have focused on communication quality and video quality such as latency and rebuffering time to estimate the QoE of a user, the QoE of a user can be affected not only by external factors that do not originate from the user, such as video quality, communication quality, viewing environment, and content itself, but also by internal factors that originate from the user, such as its current mood and personal preferences for videos. For these reasons, it is actually difficult to estimate QoE using only the quality of communication and a video. In addition, in order to use the estimated QoE for controlling video delivery, it is necessary to obtain the QoE of a user who is watching the video in a real-time manner. To solve these technical problems, we establish a method to estimate QoE using the user's biometric data. Particularly we focus on electroencephalogram (EEG), which can be obtained in real time and it is thought to be able to acquire psychological information about the user.

In our previous research, we developed a method to make a classifier for estimating QoE of a video viewing user by extracting features in the time domain and the frequency domain from the user's EEG. We clarified the estimation accuracy of the method in [3].

For realizing the QoE-based delivery control, it is important to accurately estimate the user's QoE when the video quality changes. In this thesis, in addition to our previous method, we propose a method to obtain another useful information from EEG for video delivery controls. We use event-related potentials (ERPs), which are EEGs evoked by a specific event (stimulus) and are endogenous potentials that reflect the subject's cognitive attitude toward the stimulus. The P300, which is one of ERPs, is classically known to occur in association with the presentation of low-frequency stimuli in stimulus-classification tasks such as the oddball task, but recent studies have reported that it is also evoked when perceiving changes in video quality [4]. In this thesis, we use the P300 as an indicator of the cognition of changes in video quality. By combining the P300 with the estimated QoE with our previous method, we can obtain the QoE immediately after the user's cognition of changes in video quality. In other words, we utilize the P300 as a decision criterion of the video delivery control. The overview of our proposed video delivery system is shown in Fig. 1.

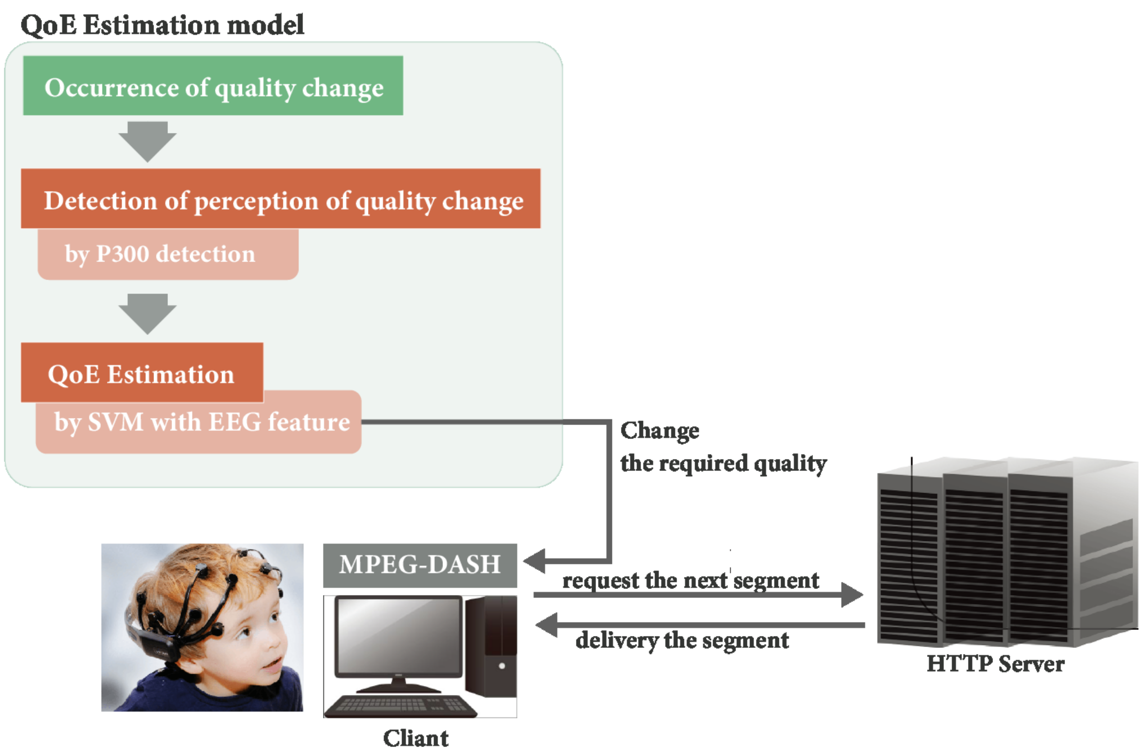

Figure 1: Overview of our proposed video delivery system.

To evaluate our proposed method, we need a data set which contains the QoE and EEG of video viewing users. Therefore, we start with data collection experiments. After obtaining a data set, we evaluate our proposed QoE estimation method and P300 detection method. Finally, to verify the feasibility of the real-time video rate control with our proposed method, we implement an MPEG-DASH player that request a bitrate determined by using the estimated QoE right after the detection of the P300, to a video server.

The remainder of this paper is organized as follows. In Section 2, we briefly describe related work about ERPs and QoE-based delivery control methods. In Section 3, we explain the experiments that we conducted to collect data set for the estimation of QoE. In Section 4, we present our QoE estimation method and show the evaluation results of our proposal. Also, in this section, we present a P300 detection method. In Section 5, we show the detailed design of the video delivery system with MPEG-DASH and a QoE-based bitrate control. Section 6 gives the conclusion of this thesis.

2 Related Work

2.1 Event Related Potential in the EEG

Event Related Potentials (ERPs) are voltage fluctuations generated in the brain elicited by stimuli or events. The P300 is one of the event-related potentials, a positive voltage fluctuation that appears near the parietal lobe about 300 ms after the perception of a significant stimulus (e.g., target stimulus in Oddball paradigm [5]). The P300 is often used in the field of Brain Computer Interface (BCI), such as the P300 Speller.

For the detection of the P300, the additive averaging process is generally applied to improve the SINR (Signal to Interference and Noise power Ratio) and obtain high reliability [6]. For single-trial detection of the P300, amplitude threshold-based methods and linear discriminant analysis-based method have been proposed [7, 8].

The P300 has also been utilized to assess the video quality of multimedia [9, 10]. In general, it is known that the amplitude of the P300 is smaller in passive task conditions, where no information about tasks and stimuli are presented to subjects in advance, than in active task conditions [11]. There is not enough discussion on whether it is possible to detect the P300 for unexpected changes in video playback such as bitrate changes and rebufferings.

2.2 QoE Estimation Methods in Video Streaming Services

QoE is a measure of the user's satisfaction with its enjoying services. QoE is affected by system factors, context factors and human factors [12]. System factors include communication delay, packet loss, video resolution and frame rate. References [13, 14] show that rebufferings in video watching have a large impact on user's QoE. References [14,15] show the played bitrate and the bitrate change ratio are also affect the QoE.

Context factors include location, time of day, and type of the tasks. References [16,17] describe that the QoE is video content dependent, and Ref. [18] discusses the utility of social network trend analysis for QoE estimation through a case study involving simulation of HTTP adaptive streaming (HAS) and load balancing in a content distribution network (CDN).

QoE also relies on Human internal factors like gender, age, mood, and cognitive processes expectations. To reflect human cognitive processes for QoE estimation, some studies use EEG data of users [4, 9, 10]. By combining these multiple factors, it is expected that QoE can be estimated more accurately and effective control can be achieved.

2.3 Adaptive Video Streaming Techniques

In recent years, HTTP Adaptive Streaming (HAS) has been widely used in video delivery services. For example, it has been implemented in Microsoft's Silverlight SmoothStreaming (MSS), Apple's HTTP Live Streaming (HLS), and Adobe Systems' Adobe HTTP Dynamic Streaming (HDS). One of the HAS standards is DASH published by MPEG in 2012 [19]. DASH aims to provide users with a smooth video streaming service based on network conditions and types of client devices. In DASH system, the video content is encoded into multiple versions at different bitrates, and each encoded video is divided into segments of a certain length. These segments are stored in the DASH server, and the server delivers the specified segment in response to a request from a DASH client. The DASH client requests segments of suitable quality according to a predefined ABR algorithm. Various ABR algorithms have been proposed, for example, throughput based control methods in [20, 21] and buffer based control methods in [22, 23].

3 Experiment for EEG Data Collection

To implement and evaluate a method for estimating the QoE of a video viewing user, we conducted an experiment to collect QoE and EEG during video viewing. 25 healthy students from our University participated in the experiment. Our experiment received approval from Osaka University Research Ethics Committee and permission from the head of our research institution. Considering the recent spread of COVID-19 infection, we carried out experiments in a ventilated room. Also, both the person conducting the experiment and the subjects checked own body temperature, disinfected their hands and fingers.

3.1 Device Configuration

Figure 2 shows the experimental environment. In the experiment, three types of data were recorded: EEG data, QoE data, and playback start time. Two laptops were used, one for video playback and the other connected to an EEG sensor (Emotiv EPOC X) for recording EEG raw data. Each laptop for video playback was connected to a 27-inch monitor and participants watched a video clip on the 27-inch monitor. Laptops and smartphones (the details will be described later) were connected to the same network through a wireless router placed in the same room, and their Operation-system time was synchronized with an NTP protocol to keep time difference below the order of 10ms. Time synchronization is sufficient for our system to detect the P300.

3.2 Experimental Protocols

Figure 3 describes the protocol of the experiment we designed. Before the start of the experiment, the time of each device was synchronized using an NTP. The video view trial was repeated 30 times, corresponding to the number of total video clips. In order to avoid participant from fatigue affecting the measurements, a 10-minute break time was provided every 10 trials.

Figure 2: Experimental environment.

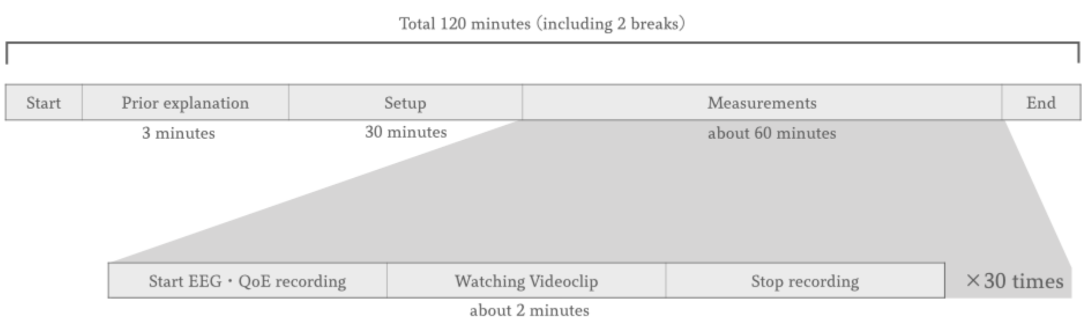

Figure 3: Protocol of experiments.

3.3 Video Clips

For our experiment, we used the video clips provided in LFOVIA database [24]. LFOVIA database contains unprocessed video clips and processed video clips. The number of unprocessed video clips is 18. 18 video clips are processed with two types of rebuffering and bitrate variation. Therefore, the number of processed video clips is 36. In our experiment, we used 30 processed video clips from 36 processed video clips. The LFOVIA database also contains time-series information of the bitrate of each video clip.

Video clips are played by a VLC media player. In order to focus only on the variation of EEG and QoE in response to the variation of video quality, subjects are prohibited to operate a seek bar and pause the video during playback. The time of the start of video playback is recorded and used for the synchronization with EEG data.

3.4 EEG measurement

EEG data is measured by Emotiv EPOC X and recorded by their software EmotivPro (Emotiv [25]). EPOC X has 14 symmetrical sensor electronodes, which are AF3–4, F7–8, F3–4, T7–8, P7–8 and O1–2 of the international 10–20 system. The resolution and the sampling rate of EPOC+ are 14 bit and 128 Hz, respectively. The recording of the EEG started before the start of playing video and stopped after the video was finished. After measurement, EEG data was trimmed in the time interval the video was playing.

Figure 4: EPOC X for recording EEG in the experiment.

3.5 QoE measurement

The Single Stimulus Continuous Quality Evaluation (SSCQE) [26] was used to evaluate the quality of the stimuli. In this method, the subject continuously scores on a scale from 0 to 100 while the video is playing. Participants entered QoE scores by manipulating the slider. The value of the slider is obtained every 0.5 s. In our experiment, we followed the SSCQE to obtain QoE of subject, but we provided smartphones to subjects for input of QoE scores. We made an application run on the smartphone to input QoE scores by the slider (Fig.5).

During the measurement, the screen of the smartphone is mirrored on the same screen as the video, by using a mirroring application, so that the subject can check the QoE value being entered at its hand. In addition, a vibrating function is used to provide feedback on the grade being entered. When the value being entered is 0 or 100, the device vibrates to inform the subject. The measured data were sent to the computer playing videos via WebAPI.

Figure 5: QoE recording application (Runs on iPhone 12)

4 QoE Estimation method

We propose a method to estimate QoE of video viewing users. We construct an SVM-based classifier. Following sections explain the methods for constructing the QoE classifier using EEG data; a preprocessing method of the EEG, a feature extraction method from the EEG, a feature selection method using a GA, and a classification method.

4.1 Preprocessing of EEG data

A Butterworth bandpass filter was applied to remove irrelevant artifacts such as electromyogram (EMG) and power supply noises from the measured EEG raw data. We used MATLAB to apply the bandpass filter to EEG raw data. For subsequent feature extraction, the bandpass filter was applied in the four bands as shown in Table 1, where Fstop1 is the endpoint frequency of the first blocking band, Fstop2 is the starting frequency of the second blocking band, Fpass1 is the starting frequency of the passband, and Fpass2 is the endpoint frequency of the passband. Data contaminated by the subject's body movements and the disruption of communication during recording were excluded.

Table 1: Bandwidth of the Butterworth bandpass filter

Bandwidth	Fstop1	Fpass1	Fpass2	Fstop2
total	3.8	4	30	31
θ	3.8	4	8	8.2
α	7.8	8	12.5	12.7
β	12.3	12.5	30.5	31.7

4.2 Feature Extraction

After removing the artifacts from the measured EEG using a bandpass filter, four kinds of features were extracted every 2-s non-overlapping hamming window. The extracted features are Band power, Power Spectral Density (PSD), and Distributed Wavelet Transform (DWT), which have been commonly used in earlier studies [9,27]. We used MATLAB to remove artifacts and also to compute these features. From PSD and level 2 to 4 components of DWT, four types of features are calculated; median, maximum, minimum and variance.And the normalized frequency at which the power of the signal is maximized was also calculated from PSD. We calculated Band Power, PSD and DWT features for each of the four frequency bands; 1) theta 4–8Hz, 2) alpha 8–12.5Hz, 3) beta 12.5–31Hz and 4) total 4–31Hz. In addition, the ratios of Band power from 1) to 3) and Band power from 4) were also calculated. We calculated each of features on the 14 channels of the EEG sensor, so the dimension of the total features is 546.

4.3 Feature Selection

We limit the number of features to be used for QoE estimation to reduce the calculation time. To achieve better accuracy even with a limited number of features, we used a GA to find a good set of features. A GA is a meta-heuristic algorithm inspired by the process of natural selection, where the optimal individual is selected by selecting highly adapted individuals from the population and repeating the mating and mutation processes. We used a GA to explore the best combination of features for training the classifier. We used DEAP library in Python to implement a GA.

Individuals of the GA in our proposal were represented as a permutation of integers from 0 to 545, where each integer corresponded to a type of feature. The fitness of an individual was defined as the average accuracy of the three-part cross-validation by an SVM with a Radial Based Function. For calculating the fitness, the features from the first to the third, fifth and tenth genes in individuals were used. The number of individuals and generations of the GA was set to 30 and 100, respectively. And the mutation and crossing probabilities of the GA were set to 0.01 and 0.6, respectively, and the mutation probability was raised to 0.5 every 100 generations to prevent convergence to a local solution. The selection function used the tournament system. In the tournament system, a certain number of individuals are randomly selected from the population and the one with the highest fitness is selected for the next generation.

The crossover function used the partial matching crossover method. Partial matching mating is a method of mating in which a gene for mating is randomly selected from two individuals and the genes are rearranged so that the sequence of the selected gene is the same as that of the other individual. The mutation function used a translocation method in which the gene arrangement within an individual was randomly switched.

4.4 QoE Classification Method using Support Vector Machine

As a baseline classification method, we used an SVM with a Radial Based Function (RBF) kernel as a classifier. The RBF Kernel has two parameters; C determines the penalty of misclassified data and γ determines the width of the RBF kernel. We used scikit-learn library in Python to implement an SVM. The kernel parameters were set to $C = 1.0, \gamma = 1/546$.

The QoE ratings collected in the experiment were used to label for learning. The labeling was done for each video. First, the QoE quartiles were calculated and the bottom 25% were labeled Bad and the top 25% were labeled Good. To determine the performance of the classifier in estimating the event of QoE decline, which is important for delivery control, we trained the classifier as a binary classification problem using only Good and Bad data. We trained the SVM for each participant and evaluate the recall of Bad label. We divided the data from the 30 trials into test data and train data evenly.

4.5 Evaluation of QoE Classification

We evaluated the performance of the QoE classifier regarding the recall for different feature selection methods.

Figure 6 shows the recalls for each subject. The maximum value is 74%, the mean value is 49.3%, and the minimum value is 19%. As shown in Fig. 6, there is a large difference in the recall values among the subjects. This is due to individual differences in QoE criteria of subjects, and it is necessary to verify the classification results when combined with other information such as content characteristics and QoE trends.

4.6 P300 Detection

To detect the user's perception of changes in video quality, we describe the method to detect the P300 from the EEG data. At first, a band-pass filter of 0.7-7 Hz was applied to the EEG Raw Data to remove frequency components not associated with the P300. Then signals of all channel were averaged across all channels, and moving average value with a

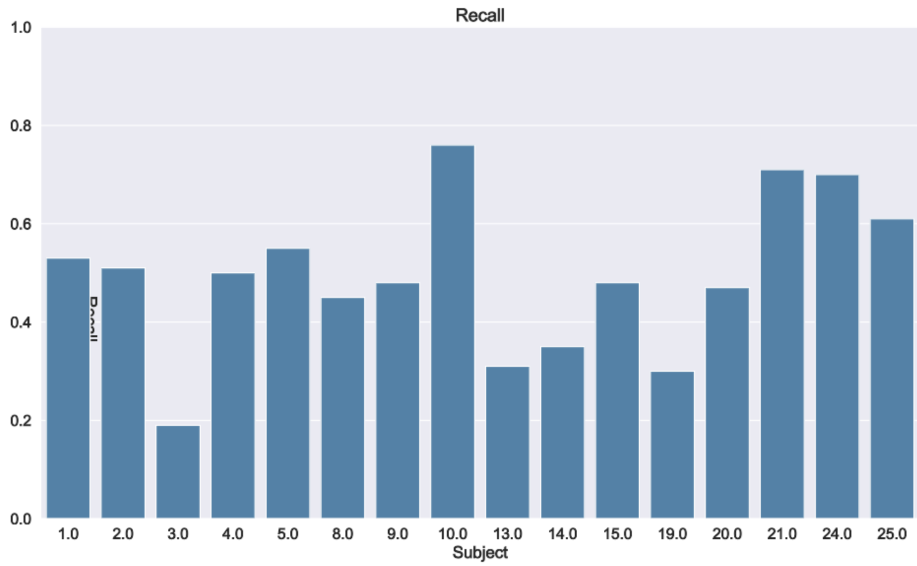

Figure 6: Recall of each subject.

window size of 20 ms was taken.

Because requesting QoE input by a smartphone might make a subject to watch the video actively, which affect the magnitude of the P300 amplitude, the following analysis used data from subjects who did not input QoE.

We adopted a threshold-based detection method for the P300. The P300 is detected when amplitude of the EEG signal exceeds the predefined threshold value [7]. Before deciding the threshold, we compared the waveforms when quality changes occurred with those when stable playback was taking place. And then, we verified whether P300 could be detected by threshold-based method.

As shown in Fig. 7, comparing the average amplitudes at the time when no video quality change occurs and the time between 200 ms and 1000 ms after the rebuffering, we can see that the amplitude is larger after the rebuffering. This result suggests that we could be able to detect the P300 in the video playback scenario.

From the baseline signal, the mean amplitude μ and standard deviation σ were calculated and they were used to calculate the threshold. The threshold λ is determined by the following equation.

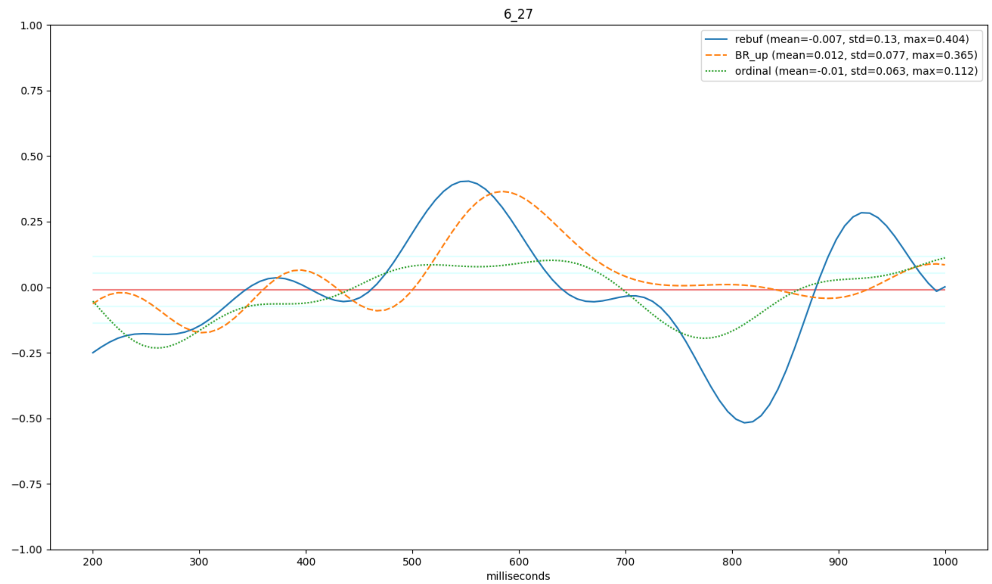

Figure 7: Waveforms compared by each event, rebuffering, bitrate changes, and baseline.

$$\lambda = \mu + a\sigma \tag{1}$$

To determine the threshold parameter a, the P300 detection was performed while varying it by 0.5 between $a \in [2, 3.5]$. And we chose $a = 3$ as it had the lowest noise.

Figure 8 shows the plot of bitrate value and the timing of P300 detection of a subject (subject number is 6) when the parameter a is set to 3. Blue line shows the bitrates and orange area shows the time period where the P300 was detected. In Fig. 8, the P300 is detected around the bitrate change in video clips 4, 14, and 28, but not detected after the bitrate change in video clips 9, 16, and 19.

In the case the P300 is not detected after video changes, this subject was either not aware of the quality variation, or was aware of the quality changes but the amplitude above the threshold has not appeared. In the former case, the content of the video clip, the frequency and duration of the quality changes might be affecting the subject. And in the latter case, the threshold might have been set higher than the appropriate value, so noise affect adversely the P300 detection. In order to identify the cause of the problem, we need to investigate the relationship between the characteristics of the video clip and the detection rate, and the change in the detection rate when the noise removal method is improved.

Figure 8: P300 Detection with QoE and Bitrate value (Subject 6).

5 QoE-based Delivery Control

5.1 Video Delivery Control based on P300 Detection and Estimated QoE

As we described in Sec. 2, The P300 is an event-related potential and is a brain response that appears as a result of human thought and cognition. And the appearance of the P300 reflects active attention of a human, which is confirmed in an experiment called the oddball experiment. The P300 has also been noted to appear in passive attention. We investigate the realization of a video delivery system that acquires the timing at which the P300 occurs due to rebuffering and rate changes that occur during video playback, and determines the video bitrate according to the QoE at that time. The overview of our assumed video delivery system using EEG is shown in Fig. 1.

Our basic idea in using QoE to determine the bitrate of a playing video is to grasp the minimum bitrate at which the user will not be dissatisfied, and to set such a bitrate to cause as little video rebuffering as possible. However, the QoE classifier in this paper has not been able to learn training data to take into account differences in video quality. Therefore, in this section, we describe a rate control algorithm that follows a simple algorithm to improve the QoE degradation caused by video rebuffering.

In order to prevent the user's QoE from decreasing due to a decrease in video quality, a high video quality is set when there is room for communication bandwidth, and the video quality is reduced when there is no room for communication bandwidth. The bitrate increases when the playback buffer length becomes longer than the threshold value (T_{buf}). Once the bitrate is increased, it will not be increased until N_{keep} segments have been downloaded. This prevents the bitrate from being reached to the maximum one instantaneously. When the video player detects the video rebuffering or that the quality of the video being played has changed, it checks whether the P300 has occurred between immediately after the rebuffering and 1,000 ms. If the P300 has not occurred, the current bitrate is kept; if the P300 has occurred, the QoE is estimated based on the subsequent EEG of T_{chk}, and if the estimated QoE is lower than a predetermined threshold, the bitrate is changed to one step lower than the current bitrate.

5.2 Implementation and Evaluation of the Video Delivery Control Method

In order to show the feasibility of rate control in real time, we implement our proposed control method into an MPEG-DASH client player. The system configuration of our proposed video delivery system using EEG is shown in Fig. 9.

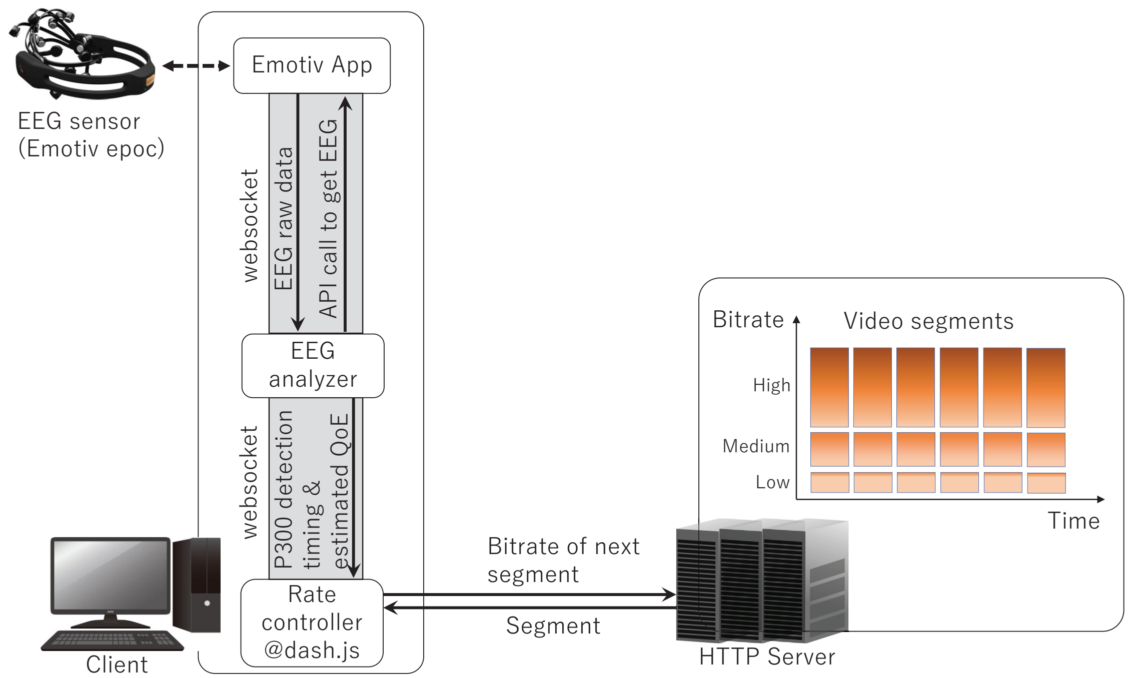

Figure 9: System configuration of our proposed video delivery system.

We build an MPEG-DASH video streaming server that controls video quality with adaptive streaming. A video client that plays a video is Chrome browser on a computer. We implemented a rate controller that determines the bitrate of the next downloaded segment in *dash.js* which is a reference client implementation for the playback of MPEG DASH via Javascript. The rate controller uses the *setQualityFor()* method in *MediaPlayer Module* to request a next segment to be downloaded with a specified bitrate to the video server.

To obtain EEG raw data, as we described in Sec. 3, we use an EPOC X device. EPOC X connects to the client computer wirelessly via a dedicated USB dongle. To access EEG raw data using an EPOC X in a real-time manner, in addition to the EmotivPRO license, we obtain the license to access Cortex SDK in *Emotiv App*. We developed an

EEG analyzer (EEG-A) in Fig. 9, which is written in python. *Emotiv App* is listening for connection via port 6868, and EEG-A connects to *Emotiv App* via websocket.

When EEG-A obtains any information from an EPOC X, the Cortex API [28] is used. The Cortex API is built on JSON and WebSockets and so it is easy to access EEG raw data from EEG-A. EEG-A conducts QoE estimation and the P300 detection method described in Sec. 4 using EEG raw data received from *Emotiv App*. Also, EEG-A connects to dash.js through a websocket on port 81 and it sends the last time of the P300 detection and estimated QoE to dash.js.

In dash.js, the *on()* method in the *MediaPlayer Module* is used to listen for public events found in the DASH player. Here, we implemeted an eventlistner to detect a rebuffering in the player by check an event of *'MediaPlayer.events["BUFFER_EMPTY"].'* If the P300 is detected within 1,000 ms after the rebuffering event, dash.js requests a segment with a lower bitrate.

The length of the playback buffer is checked every after a segment has been downloaded. If the length is larger than the predefined threshold, and if the bitrate of the last N_{keep} segments except for the last one is equal to or more than the bit rate of the last segment, then dash.js requests a segment with a higher bitrate.

Figure 10 shows a screenshot of the MPEG-DASH player using dash.js in which we implemented our rate control algorithm. we used the Big Buck Bunny movie [29], which is free to use and available in many formats. To verify the operation of our rate control algorithm, EEG-A always sent a P300 detection message. Therefore, dash.js receives the P300 detection message right after it detects the rebuffering. As a result, dash.js requests the server for a bitrate that is one step lower than the currently selected bitrate. This is confirmed by the following message displayed in the Console of the Chrome browser as shown in Fig. 10.

rebuffering happend Receive quality-change request due to P300 detection Next index of bitrate: 2

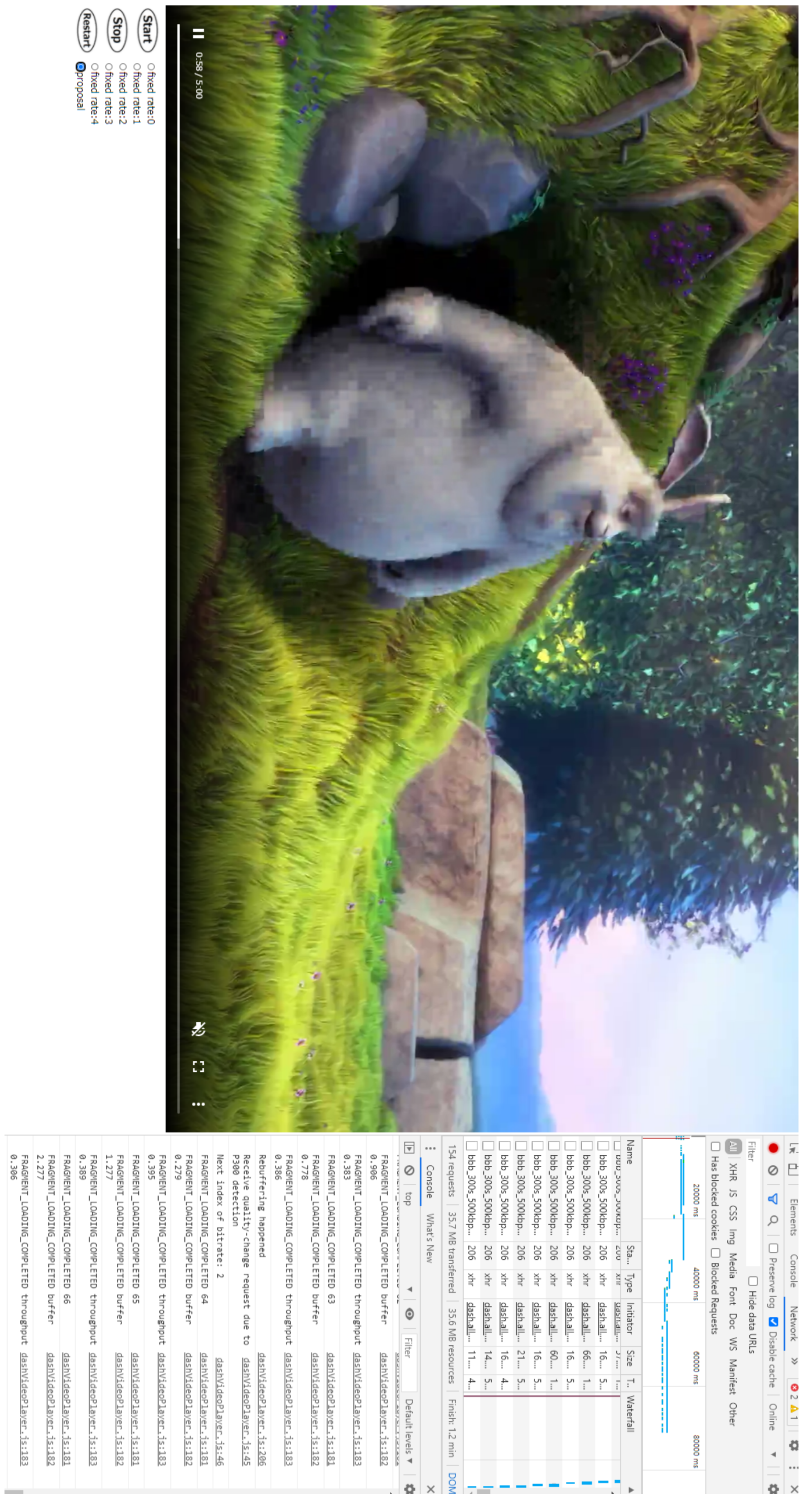

Figure 10: Screenshot of our MPEG-DASH player.

6 Conclusion

Estimating the QoE of video viewing users is an important issue for video application service providers. In this thesis, we propose a delivery control method with QoE and QoE estimation method from users' EEG data for delivery control system. Using the subject's EEG data during video viewing and QoE responses during video viewing, we trained the QoE classification model and determine P300 detection threshold. We evaluated the value of recall when bad label is classified as target, and showed that the recall value was 76.0% at maximun and the average is 49.3%. We calculated the P300 threshold from the mean and standard deviation and applied it to each subject's trial. In this case, the subject is either not aware of the quality variation, or he/she is aware of the quality variation but the amplitude above the threshold has not appeared. In the former case, the content of the video clip, the frequency and duration of the quality variation may be affecting the subject, and in the latter case, the threshold may have been set higher than the appropriate value due to noise. In order to identify the cause of the problem, we need to investigate the relationship between the characteristics of the video clip and the detection rate, and the change in the detection rate when the noise removal method is improved. Finally we considered the design of a video delivery system that acquires the timing of P300 due to rebuffering and rate changes that occur during video playback, and determines the video rate according to the QoE at that time, and implemented and verified the operation of an MPEG-dash cliant player. As a futurework, we need to investigate the accuracy of multimodal QoE classification that incorporates other information besides EEG, such as eye movements, facial expressions, and video content. the relationship between P300 detection rate and video clip features also needs further discussion. And we will construct a rate control method considering user's individual QoE and evaluate the effectiveness of the control.

Acknowledgments

There are so many people to thank for helping me during my master's degree studies at Osaka University. I would first like to express my appreciation to my thesis supervisor Professor Masayuki Murata of Osaka University. His continuous support and expensive advice throughout my studies steered me in the right direction. Engaging the tutelage of him is the valuable experience that I have ever had. I would like to thank Assistant Professor Daichi Kominami of Osaka University, for his elaborated guidance and invaluable firsthand advice. I am grateful to Associate Professor Shin'ichi Arakawa and Associate Professor Yuichi Oshita, Specially Appointed Assistant Professor Tatsuya Otoshi of Osaka University for providing me many valuable comments. In addition, I would like to thank all the members of the laboratory for support and meaningful discussion about my research and hearty encouragement. And last but not the least, I would like to express my deepest gratitude to my family for their wonderful support.

References

[1] C. Alberti, D. Renzi, C. Timmerer, C. Mueller, S. Lederer, S. Battista, and M. Mattavelli, "Automated QoE evaluation of dynamic adaptive streaming over HTTP," in *Proceedings of International Workshop on Quality of Multimedia Experience (QoMEX)*, July 2013, pp. 58–63.

[2] A. Bentaleb, A. C. Begen, R. Zimmermann, and S. Harous, "SDNHAS: An SDN-enabled architecture to optimize QoE in HTTP adaptive streaming," *IEEE Transactions on Multimedia*, vol. 19, no. 10, pp. 2136–2151, October 2017.

[3] K. Kitao, D. Kominami, and M. Murata, "GA-based feature selection for QoE estimation using EEG during video viewing," presented at 2020 International Conference on Emerging Technologies for Communications (ICETC), December. 2020.

[4] X. Liu, X. Tao, and Y. Zhan, "Calibrating human perception threshold of video distortion using EEG," in *Proceedings of IEEE International Conference on Image Processing (ICIP)*, October 2018, pp. 3543–3547.

[5] T. W. Picton, "The P300 wave of the human event-related potential," *Journal of clinical neurophysiology*, vol. 9, no. 4, pp. 456–479, October 1992.

[6] "Human auditory and somatosensory event-related potentials: effects of response condition and age," *Electroencephalography and Clinical Neurophysiology*, vol. 66, no. 4, pp. 409–419, 1987. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0

[7] B. H. Jansen, A. Allam, P. Kota, K. Lachance, A. Osho, and K. Sundaresan, "An exploratory study of factors affecting single trial P300 detection," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 975–978, June 2004.

[8] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. MÃ¼ller, "Single-trial analysis and classification of ERP components – A tutorial," *NeuroImage*, vol. 56, no. 2, pp. 814–825, 2011, multivariate Decoding and Brain Reading. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811910009067

[9] S. Scholler, S. Bosse, M. S. Treder, B. Blankertz, G. Curio, K. Muller, and T. Wiegand, "Toward a direct measure of video quality perception using EEG," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2619–2629, May 2012.

[10] S. Arndt, J. Antons, R. Schleicher, S. MÃ¶ller, and G. Curio, "Perception of low-quality videos analyzed by means of electroencephalography," in *Proceedings of International Workshop on Quality of Multimedia Experience*, July 2012, pp. 284–289.

[11] J. Y. Bennington and J. Polich, "Comparison of P300 from passive and active tasks for auditory and visual stimuli," *International Journal of Psychophysiology*, vol. 34, no. 2, pp. 171–177, 1999. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167876099000707

[12] U. Reiter, K. Brunnström, K. De Moor, M.-C. Larabi, M. Pereira, A. Pinheiro, J. You, and A. Zgank, "Factors influencing quality of experience," in *Quality of experience*. Springer, March 2014, pp. 55–72.

[13] R. K. Mok, E. W. Chan, X. Luo, and R. K. Chang, "Inferring the QoE of HTTP video streaming from user-viewing activities," in *Proceedings of the first ACM SIGCOMM workshop on Measurements up the stack*, August 2011, pp. 31–36.

[14] H. Nam, K. Kim, and H. Schulzrinne, "QoE matters more than QoS: Why people stop watching cat videos," in *Proceedings of the 35th Annual IEEE International Conference on Computer Communications (INFOCOM)*, April 2016, pp. 1–9.

[15] F. Dobrian, V. Sekar, A. Awan, I. Stoica, D. Joseph, A. Ganjam, J. Zhan, and H. Zhang, "Understanding the impact of video quality on user engagement," vol. 56, October 2011, pp. 362–373.

[16] M. Alreshoodi and J. Woods, "QoE prediction model based on fuzzy logic system for different video contents," in *Proceedings of European Modelling Symposium*, March 2013, pp. 635–639.

[17] A. Khan, L. Sun, and E. Ifeachor, "QoE prediction model and its application in video quality adaptation over UMTS networks," *IEEE Transactions on Multimedia*, vol. 14, no. 2, pp. 431–442, November 2012.

[18] T. HoÃ feld, L. Skorin-Kapov, Y. Haddad, P. Pocta, V. A. Siris, A. Zgank, and H. Melvin, "Can context monitoring improve QoE? a case study of video flash crowds in the internet of services," in *in Proceedings of the 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, July 2015, pp. 1274–1277.

[19] ISO/IEC 23009-1:2019, "Information technology   dynamic adaptive streaming over HTTP (DASH)   part 1: Media presentation description and segment formats," December 2019, accessed on January 30, 2021.

[20] C. Liu, I. Bouazizi, and M. Gabbouj, "Rate Adaptation for Adaptive HTTP Streaming," in *Proceedings of the Second Annual ACM Conference on Multimedia Systems (MMSys)*, February 2011, pp. 169–â 174.

[21] J. Jiang, V. Sekar, and H. Zhang, "Improving fairness, efficiency, and stability in HTTP-based adaptive video streaming with festive," in *Proceedings of the 8th International Conference on Emerging Networking Experiments and Technologies (CoNext)*, December 2012, pp. 97–108.

[22] T.-Y. Huang, R. Johari, N. McKeown, M. Trunnell, and M. Watson, "A buffer-based approach to rate adaptation: Evidence from a large video streaming service," *SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 187–198, August 2014.

[23] K. Spiteri, R. Urgaonkar, and R. K. Sitaraman, "BOLA: Near-optimal bitrate adaptation for online videos," *IEEE/ACM Transactions on Networking*, vol. 28, no. 4, pp. 1698–1711, June 2020.

[24] N. Eswara, K. Manasa, A. Kommineni, S. Chakraborty, H. P. Sethuram, K. Kuchi, A. Kumar, and S. S. Channappayya, "A continuous qoe evaluation framework for video streaming over HTTP," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 11, pp. 3236–3250, 2018.

[25] Emotiv©, "Emotiv," https://www.emotiv.com/.

[26] ITU-R BT.500-11, "Methodology for the subjective assessment of the quality of television pictures," 2002, international Telecommunication Union.

[27] M. R. Islam and M. Ahmad, “Wavelet analysis based classification of emotion from EEG signal,” in *Proceedings of International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Feb. 2019, pp. 1–6.

[28] “Cortex API,” https://emotiv.gitbook.io/cortex-api/ Accessed on February 1, 2021.

[29] “Big Buck Bunny,” https://peach.blender.org/ Accessed on Feburuary 1, 2021.