A Study on QoE Estimation from Heart Rate Variability Using Machine Learning

K. OTA, T. HIRAGURI, and H. YOSHINO, *Member, IEEE*
Faculty of Engineering, Nippon Institute of Technology

***Abstract*-- Quality of Experience (QoE) depends on user's psychological condition. Hence, in order to estimate QoE, it is necessary to investigate the relationship between QoE and biological information. This research investigates whether QoE, when users watched audiovisual contents, can be estimated from the heart rate variability. Feature quantities used for estimation of QoE are spectral flux, spectral centroid, the average distance of Lorenz plot representing the variability of R-R interval and the area of ellipse approximating the Lorenz plot representing the magnitude of RRI. Hidden Markov model and recurrent neural network are used for estimation of QoE from the parameters of heart rate variability. From the results estimated by these methods, feature quantities and methods required for QoE estimation are discussed.**

I. Introduction

Recently, various types of communication services have been provided due to the development of communication technology. QoS (Quality of Service) has been used for evaluating the quality of communication services. However, the quality of communication services which users feel may be different depending on the psychological condition of users even if QoS of the communication service is same. Hence, QoE has become also important for evaluating the quality of communication services.

This research focuses on streaming video service as a communication service and investigates the estimation technique of QoE of an offered streaming video service based on biological information related to stress.

Human stress conditions are known to relate to the autonomic nervous system. As a means of estimating how much the autonomic nervous system is affected by stress, there is a method using salivary amylase activity value. As another means, there are some methods using heart-rate variability[1] and electrical resistance of skin surface[2]. These methods are non-invasive methods. However, feature quantities and methods for estimation of QoE are not clarified by the previous works. Hence, this research focuses on heart-rate variability and clarifies the feature quantity and methods for realizing the QoE estimation.

II. Measurement and Analysis of Heart Rate

A. Measurement of heart rate

Electrocardiogram (ECG) is a periodic wave which has some peaks called P, Q, R, S, and T-wave, respectively. From these peaks, only R-wave is extracted. R-R interval (RRI) is obtained from the time difference between the peaks of adjacent R-waves. RRI is always fluctuating and is affected by the function of the autonomic nervous system.

This research employs a wearable heartbeat sensor myBeat (WHS-1) manufactured by Union Tool Co., Ltd. in order to obtain ECG. Heartbeat sensor is fixed to chest by a dedicated electrode band. The placement position is around the epigastrium, but before the start of experiment, appropriate position is determined by confirming the ECG and the RRI.

Since RRI is constantly fluctuating, it is sampled at unequal intervals. However, the data is required to be sampled at equal interval in order to carry out the frequency analysis. Hence, the measured RRI is interpolated with a cubic spline function and is resampled at 10 [Hz].

B. Analysis of heart rate

STFT(Short-term Fourier Transform) is carried out against the interpolated RRI. Spectral centroid and spectral flux of t-th frame are calculated by

$$C_t = \frac{\sum_{k=1}^{K} kX(t,k)}{\sum_{k=1}^{K} X(t,k)} \tag{1}$$

and

$$F_t = \sum_{k=1}^{K} \{X(t,k) - X(t-1,k)\}^2 \tag{2}$$

respectively[3]. Here, $X(t,k)$ represents the power spectum of t-th time frame and k-th frequency bin.

The Lorenz plot[4] is obtained by plotting RRI on a plane with RRI at time t as x-coordinate and RRI at time t+1 as y-coordinate. In order to analyze Lorenz plot quantitatively, Lorenz plot is approximated by an ellipse. The major axis of the ellipse is assigned to $y = x$, and the minor axis of the ellipse is assigned to a perpendicular straight-line to $y = x$. The center of the ellipse is determined by the average distance of all plots projected onto the major axis. This average distance is referred as d_t. The length of major axis is determined by the standard deviation of plots projected onto the major axis. This standard deviation is referred as σ_x. The length of the minor axis is determined by the standard deviation of plots projected onto the minor axis. This standard deviation is referred as σ_{-x}. The area of approximated ellipse is represented by

$$S_t = \pi\sigma_x\sigma_{-x}. \tag{3}$$

III. Estimation Method

A. Hidden Markov Model (HMM)[5]

HMM is a probability model which can handle the time series data. HMM defines a joint probability $p(x, c)$ against a feature quantity sequence x=<x_0, x_1, ..., x_{T-1}> and a state sequence c=<c_0, c_1, ..., c_{T-1}>. Estimation of a hidden state sequence using HMM is carried out by

$$\hat{c} = \underset{c}{\operatorname{argmax}}\, p(c \mid x) = \underset{c}{\operatorname{argmax}} \frac{p(x, c)}{p(x)}$$
$$= \underset{c}{\operatorname{argmax}}\, p(x, c)$$
$$= \underset{c}{\operatorname{argmax}}\, p(c_0) p(x_0 \mid c_0) \prod_{t=1}^{T-1} p(x_t \mid c_t) p(c_t \mid c_{t-1}) \cdot \quad (4)$$

In this research, each state c_t corresponds to one of the discrete integer values from 1 to 5. A feature vector consists of $(C_t, F_t, d_t, S_t, C_t - C_{t-1}, F_t - F_{t-1}, d_t - d_{t-1}, S_t - S_{t-1})^{\mathrm{T}}$. Each feature quantity x_t corredponds to one of cluster numbers from 1 to N. Cluster numbers can be obtained by clustering the feature vectors by k-means algorithm. State transition model is shown in Fig. 1. Only transition to the adjacent nodes is allowed in this model.

B. Recurrent Neural Network(RNN)[6]

RNN is the neural network with a feedback loop. RNN can take consideration with the history of input data in calculating the current output owing to this feedback loop.

However, it is said that the fundamental RNN cannot consider the long-term history of input data. In order to overcome this defect, LSTM (Long Short-Term Memory) is introduced to the fundamental RNN. The structure of LSTM is realized by replacing each hidden node of the fundamental RNN into a memory unit. Memory unit has a forgetting gate which controls the degree of forgetting.

IV. Subjective Evaluation of QoE

Subjective evaluation is carried out in an anechoic chamber in Nippon Institute of Technology. In this experiment, subjects are required to evaluate QoE of audiovisual contents whose resolution is degraded and/or that include 15-seconds halt. The resolution is set to 1280 × 720 (original size) or 854 × 480. Subjects are required to watch a total of 8 kinds of audiovisual contents. During the video viewing, each subject constantly evaluates the QoE for the audiovisual contents by the tablet in 5 levels (Bad “1”, Poor “2”, Fair “3”, Good “4”, Excellent “5”) with wearing the heartbeat sensor. The evaluation value is sampled at 10 [Hz].

V. Estimation Results

Training for HMM and RNN is carried out by QoE and feature quantities of heartbeat obtained during viewing seven out of eight audiovisual contents and the performance evaluation is carried out by the rest of data. Table I shows the parameter setting for HMM and RNN. Performance is evaluated by the concordance rate and correlation coefficient between measured and estimated QoE. In this research, RNN estimates QoE of continuous value, so estimated QoE whose error is within 0.05 is considered as coincident with measured QoE.

Average concordance rate of HMM over eight audiovisual contents was 0.356 and average correlation coefficient of HMM was 0.129. Average concordance rate of RNN over eight audiovisual contents was 0.269 and average correlation coefficient of HMM was 0.12. An example of estimation results shows in Fig. 2.

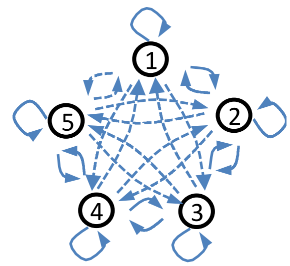

Fig. 1 State transition model. Transition probability shown by the dashed arrow is nearly zero.

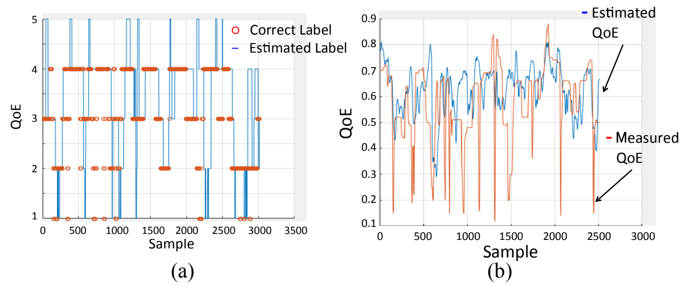

Fig. 2 An example of estimation results of (a) HMM and (b) RNN with LSTM. In case of HMM, concordance rate is 0.497 and correlation coefficient is 0.371. In case of RNN, concordance rate is 0.3 and correlation coefficient is 0.196.

Table I Parameters of HMM and RNN.

Parameters	Values
# of clusters (HMM)	100
# of nodes of input layer (RNN)	8
# of nodes of hidden layer #1 (RNN)	160
# of nodes of hidden layer #2 (RNN)	80
# of nodes of output layer (RNN)	1
Forgetting factor (RNN)	1.0
Dropout (RNN)	0.6

VI. Conclusion

This research focused on heart-rate variability in order to estimate QoE. As a result, About 30 % - 40 % QoE can be estimated by HMM and RNN. However, the estimation performance was not enough high. Hence, it is necessary to increase the number of feature quantity and to use the biological information except heartbeat. Moreover, it is required to investigate the relationship between QoE and emotion.

Reference

[1] H. Minamitani et al., "Heart Rate Variability related with Mental Stress and Cumulative Fatigue - Assessment of the Autonomic Nervous Functions by Using Neural Networks -," *IEICE Technical Report*, MBE98-156, pp. 115-122, Mar. 1999.

[2] R. Yamamoto, and T. Yamazaki, "Study of Quality Evaluation Using an Objective Evaluation Method for Video Streaming Services, " *IEICE Technical Report*, CQ2016-104, no. 403, pp. 81-86, Jan. 2017.

[3] M. Okemoto and M. Miura, "Estimation of perceived emotion from heart rate variability, " *MA Technical Report*, pp. 43-48, Dec., 2017.

[4] N. Terasawa, and N. Bu, "Evaluation of Visually Induced Motion Sickness based on Heart Rate Variability using Lorenz Plot, " *IPSJ SIG Technical Report*, pp. 1-6, 2015.

[5] L. Rabiner, "An Introduction to hidden Markov models", *IEEE ASSP Magazine*, pp. 4-16, 1986.

[6] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling", *In INTERSPEECH-2014*, pp. 338-342, 2014.