Detection of Media Enjoyment using Single-Channel EEG

Zhen Liang*, Hongtao Liu*, and Joseph N. Mak*

***Abstract*—We propose to use real-time single-channel EEG signal to quantify user enjoyment level elicited by media content. Selected time-frequency components from single-channel frontal EEG were extracted and formed a statistical multivariate model predicting user enjoyment level. Frequency components from Theta, Alpha and Beta bands at different time moments were selected. We found robust model performance during 10-folds cross-validation with 100 repetitions. A high correlation of around 0.8 between predicted and actual enjoyment level of subjects was achieved. Considering various factors of the selected features, we found an important role of alpha as the emotional component, and beta as the cognitive component involved in the complex enjoyment processes. Also, in accordance with the peak-end rule, feature from latter part of the video seems to create a large influence to the overall experience. From all of these results, we implement real-time EEG-based detection system for media user enjoyment with single EEG channel.**

***Keywords*— Electroencephalography (EEG); media content; enjoyment; appreciation; forehead.**

I. INTRODUCTION

Enjoyment of media content or entertainment is a complex phenomenon. As suggested by disposition-based theories [1-3], the media enjoyment process centers on different factors, including user's emotional responses, cognition, personal experiences, values and favor. The earliest disposition literature utilized the term ***appreciation***, and later ***enjoyment*** in last decade, when referring to one's hedonic response to media content [3]. Regardless of the nature or form of media content, people, in everyday speech, use the term *enjoyed* to describe their experiences while listening to pleasant music, viewing horror movies, or other different types of media content.

Emotions play a critical role in how people select and appreciate any media content. Seeking to find "what user like" and understand "why user like" are goals shared by psychologists, media content providers and tech-companies. Digital media collections are growing rapidly with an increasing capacity and content variety. Multimedia is supposed to be entertaining and made to induce enjoyment. Emotional features of multimedia are therefore an invaluable source of information for multimedia indexing and recommendation [4, 5]. Given the individual nature of enjoyment process, the availability of methodologies for automatically extracting this type of information will make possible the advancement of media content analysis, affective indexing and retrieval technology.

Comparing to other affective computing or emotion recognition methods detecting facial expressions and speech, methods based on physiological signals measurement, such as skin conductance, Electrocardiogram (ECG), and Electroencephalogram (EEG), provide a continuous and objective measure of the user's inherent responses to media content. Amongst these physiological signals, EEG provides a direct measure of the brain, which is the primary generator of mental activities relating to media enjoyment.

Different important works in emotion detection using EEG have been reported in the past decades. One of most important findings in this area would be the frontal alpha-asymmetry, which showed repeatedly to own a strong relation with emotion, motivation and reactions to perceived stimuli [6-8]. Most of these works have focused on detecting physiological response activity related to dimensional changes in valence and arousal space [9], paying less attention to the conscious experience of an emotion. As stated above, the media enjoyment process involves a complex interplay between emotional responses and cognition, taking into account also other personal factors. Limited work has been done in exploring the brain activities behind this complex enjoyment processes which comprise not only emotional reactivity to stimuli, but also people's cognitive interaction with emotional responses generated. Moreover, most if not all of the current studies employed conventional multi-channel EEG recording system, which is not practical to be used in daily life by the general public. This becomes one of the biggest hurdles of wide dissemination of affective computing technologies.

In light of this, this study set out to investigate the neural correlates of media enjoyment using single-channel EEG signal collected from forehead region. In addition, we attempted to build a predictive model which is capable to provide a direct, intuitive measurement of user enjoyment level in form of linear scale, through single channel EEG.

II. EXPERIMENTAL DESIGN AND METHODS

A. Experiments

Seven healthy volunteers (age: 28±3 years; male/female: 4/3) were recruited in this study. All volunteers are right-handed and have normal or corrected-to-normal vision. They are naive subjects who have no experience in similar experiment and have no prior knowledge about the study. They have no history of neurological or psychiatric disorders

*All authors contributed equally to this work.

Zhen Liang is with NeuroSky, Inc. Unit 317, Lakeside One, Phase 2, Hong Kong Science Park, Shatin, New Territories, Hong Kong (phone: 852-26074337; fax: 852-39545577; e-mail: jliang@ neurosky.com).

Hongtao Liu is with NeuroSky, Inc. Unit 317, Lakeside One, Phase 2, Hong Kong Science Park, Shatin, New Territories, Hong Kong (e-mail: hliu@ neurosky.com).

Joseph N. Mak is with NeuroSky, Inc. Unit 317, Lakeside One, Phase 2, Hong Kong Science Park, Shatin, New Territories, Hong Kong (e-mail: jmak@ neurosky.com).

and are not in use of any recreational drugs or other medications.

Each subject was asked to participate in 6 experimental trials. Each trial is composed of two sessions: Baseline Collection and Video Watching. Subject EEG at a relaxed state without video stimuli was collected as baseline, while EEG during video watching was collected as task data. Audiovisual stimuli from videos involve six emotions (joy, sad, fear, disgusting, love and anger) were presented separately to subjects. During the stimuli selection, experimenter checked and confirmed each stimulus was dominated by one emotion only. After each trial, participants were asked to rate their overall enjoyment level during video watching as:

Level 1: not enjoy at all;

Level 2: low enjoyment;

Level 3: medium enjoyment;

Level 4: high enjoyment.

B. Data Collection

Subjects were asked to seat in a comfortable chair with their mind relaxed and eyes open. Experimenter would give each subject an introduction of the experiment in the beginning. Each trial last around 10 min. Sequence of videos was counterbalanced with a separation of at least 2 hours between each trial to avoid any possible priming effect.

A 30 seconds baseline was collected in the beginning of each trial. Then, 1 minute video stimulus was presented to the subject with a 24’’ LCD display monitor positioned 60cm from the subject’s forehead. Subjects were encouraged to stay focus and avoid any unnecessary movements during the trial (Fig. 1).

During the experiment, EEG data were recorded simultaneously using a single-channel dry sensor mobile EEG system (NeuroSky MindWave Mobile headset) at a sampling rate of 512Hz. Then the collected raw EEG signal were transmitted wirelessly and stored to the computer during the experiment.

Figure 1. EEG data collection during video watching.

C. Data Processing and Analysis

Raw EEG signal was first detrended. Wavelet-based filtering method [10] was then applied to the data to remove blink and eye movement related artifacts. Short Time Fourier Transform (STFT) is a well known technique in signal processing to analyze non-stationary signals, like EEG. It truncates the whole signal into short data frames and then multiplies a window so that the edified signal is zero outside the data frame. In this study, STFT was employed to analyze EEG signal.

To reduce the between-subject variance, STFT feature results were normalized by subtracting the baseline data from the video task data as follow:

$$F_N^i = F_V^i - \bar{F}_B^i$$

where F_V^i was the extracted STFT results in video task in ith trial, and $\bar{F}_B^i$ was the average power spectra in each frequency bin during baseline.

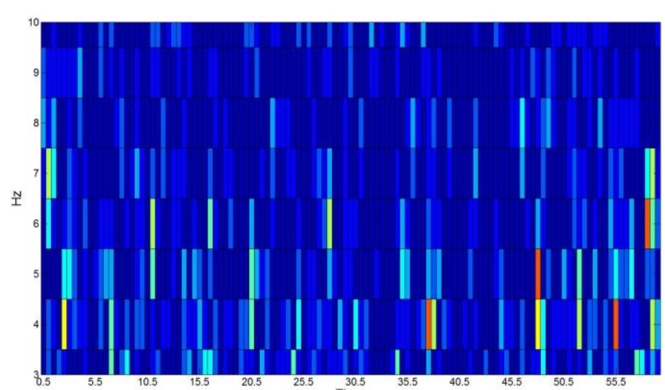
Dimension reduction was performed to the acquired feature map using Stepwise multivariate linear regression analysis [11]. Features were selected in a stepwise manner to remove irrelevant or redundant characteristics and retain the relevant features until no further improvement in the model. In the iterative process, feature candidates in the feature space were added and removed in a multi-linear model and the statistical significance in the regression model were evaluated. The statistical significance is measured by one-way analysis of F-statistic.

Further analysis of selected features was performed after the reduction of feature space size. We then explored the relationship between retained features with participants’ self-reported enjoyment level, and built a statistical multivariate regression model with the features that achieved the best performance. More details about the parameter settings in the methods and the corresponding results would be described further in Section 3.

III. RESULTS

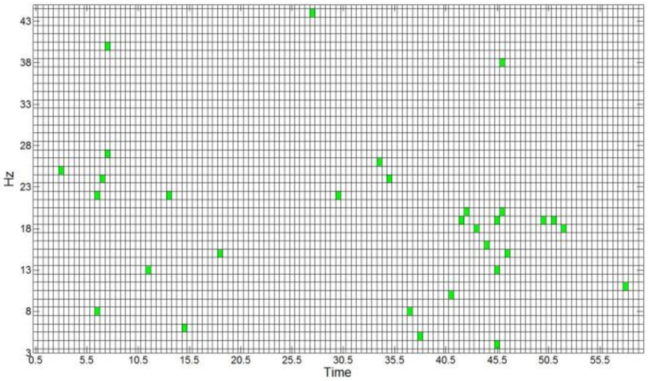
A total of 42 trials (7 subjects x 6 trials) were collected in this study. STFT features with a time resolution of 0.5 second were extracted with the preprocessed data with a Hamming window of length 512 and a 50% overlap. Focusing on frequency range of 3Hz to 45Hz, feature matrix of size 43×59 and 43×119 were extracted from baseline (F_B) and task data (F_V) respectively. Fig. 2 depicted the spectral power distribution of a subject along the time during video watching.

Figure 2. Time-frequency distribution during video watching (X axis is time, and Y axis is frequency).



The stepwise multivariate linear regression analysis with a leave-one-out replication was then applied to the calculated feature after normalization. Statistical criteria for including or rejecting a variable in a model was set to be 0.05 and 0.10 respectively. 31 features significantly improving ($p<0.05$) the model's ability to predict participants' self-reported enjoyment level were selected (Fig. 3).

Figure 3. Selected STFT features highlighted in green. X axis indicates video time from 0.5s to 60s and y axis represents spectral frequency from 3Hz to 45Hz.



Min-max normalization with range of 0 to 1 was applied to the 31 retained features. Exhaustive search of best features combination was then conducted to find the best predictive model for user enjoyment. A multivariate linear regression model was built with five selected frequency components from theta, alpha and beta at different time moments:

$$Enjoyment\ level = 0.574 - 1.951 \times F_1 - 1.092 \times F_2 + 1.944 \times F_3 + 2.511 \times F_4 + 1.513 \times F_5 \quad (1)$$

TABLE 1. SELECTED FEATURES IN THE PREDICTED MODEL OF USER ENJOYMENT (* BIVARIATE CORRELATION STATISTICAL SIGNIFICANCE AT P<0.01; ** P<0.001)

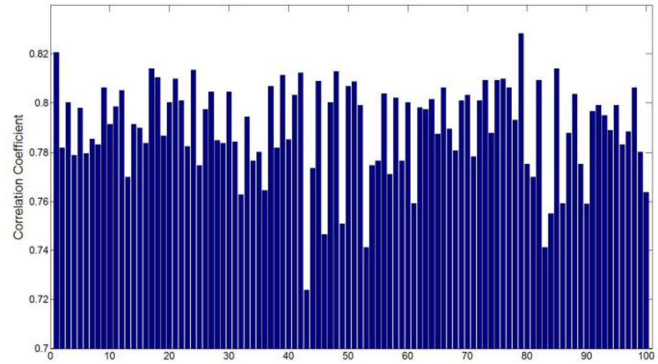
	F_1	F_2	F_3	F_4	F_5
Hz	25	24	13	16	4
Time (s)	3	7	11.5	44.5	45.5
Correlation (r)	-0.03	-0.17	0.25	0.63**	0.30*

To verify the robustness of the predictive model shown in Eq. (1), ten-fold cross-validation of testing data was performed with 100 repetitions. As shown in Fig. 4, testing results achieved significant correlation with subjective enjoyment level with an average correlation value of 0.7897 ± 0.0177. We also tested the model's performance on individual subject level. Table 2 reported a high correlation of predicted enjoyment level from model and user subjective rating in most of the subjects.

TABLE 2. MODEL PERFORMANCE FOR INDIVIDUAL SUBJECT (*STATISTICAL SIGNIFICANCE AT P<0.01; ** STATISTICAL SIGNIFICANCE AT P<0.05)

	S1	S2	S3	S4	S5	S6	S7
Correlation (r)	0.92 **	0.83	0.62	0.93 **	0.82 *	0.95 **	0.96 **

Figure 4. Correlation coefficients in 100 times 10-cv.



IV. DISCUSSION

In this study, we developed a predictive model of media enjoyment level using single channel EEG. Frequency components from Theta, Alpha and Beta bands at different time moments were selected to form the statistical model.

Alpha asymmetry was known to be a strong finding in emotion research in the past decades [6-8]. Other than alpha-asymmetry, other power spectral changes such as alpha power, theta power, frontal midline theta power, beta power, and gamma power have also been reported in emotion detection research [12-15].

These previous results seem to be different from our finding (Table 1) in this work. This could probably due to the difference of enjoyment process and pure emotion reactivity, and the spectral power decomposition in time moments in this study. As suggested by disposition-based theories, complex affect-cognition interaction happens during the media enjoyment process.

Alpha component at 13 Hz was included in the predictive model of media enjoyment. Despite its importance role in emotion detection reported by previous research, alpha component does not have the strongest predictive power in the model. Beta at 16Hz played the most important role in the predictive model of media enjoyment reported in this work. Despite being less highlighted than alpha power in emotion research, few studies reported an increase in beta power following an unspecific increase of emotional arousal that was independent with valence [16, 17]. Relationship between EEG beta power and attention and other cognitive processes were reported [18]. The significant role of beta in the enjoyment model suggests a major involvement of cognitive processing in the media enjoyment process.

Theta included as one of the predictive feature in the current enjoyment model has been reported to be an indicator of both cognitive and emotion activities [19, 20]. Complex interplay between emotion and cognition involvement in the enjoyment processes makes the distinction of these components difficult.

Single-channel EEG data was decomposed into feature matrix through STFT, making investigation of timely dynamics of spectral modulation possible. Regardless of emotion types, strong predictive features were found mainly around the 5^{th}, 10^{th}, and 45^{th} second in the 1 minute video.

From the feature weight in the multivariate model (Eq. 1) and also the bivariate correlation between features and subjective enjoyment (Table 1), we can see a stronger influence of overall enjoyment rating by the latter part of the video. This finding is consistent with what Peak-end rule proposed by Fredrickson and Kahneman [21, 22].

In sum, the result of this work reported the EEG activities behind the complex processes behind media enjoyment. Also, we demonstrated the possibility of detecting user media enjoyment level using single-channel EEG. The application of this finding leads to huge potential in the development of novel affective computing technology with digital media. Such human-centric interaction technology has the potential to revolutionize digital entertainment, learning, and many other areas of life and its need is every growing.

REFERENCES

[1] D. Zillmann and J.R. Cantor, "Directionality of transitory dominance as a communication variable affecting humor appreciation," *Journal of Personality and Social Psychology*, vol. 24, no. 2, 1972, pp. 191-198.

[2] A. Raney, "Expanding disposition theory: reconsidering character liking, moral evaluations and enjoyment," *Communication Theory*, vol. 14, no. 4, 2004, pp. 348-369.

[3] A.A. Raney, "The psychology of disposition-based theories of media enjoyment," in: *Psychology of entertainment*, Bryant, Jennings (Ed); Vorderer, Peter (Ed), 2006, pp. 137-150.

[4] A. Hanjalic and L.Q. Xu, "Affective video content representation and modeling," *IEEE Transactions on Multimedia*, vol. 7, no. 1, 2005, pp. 143-153.

[5] M. Soleymani, S. Asghari-Esfeden, M. Pantic, and Y. Fu, "Continuous emotion detection using EEG signals and facial expressions," *IEEE International Conference on Multimedia and Expo (ICME)*, 2014, pp. 1-6.

[6] R.E. Wheeler, R.J. Davidson, and A.J. Tomarken, "Frontal brain asymmetry and emotional reactivity: A biological substrate of affective style," *Psychophysiology*, vol. 30, no. 1, 1993, pp. 82–89.

[7] J.A. Coan and J.J. Allen, "Frontal EEG asymmetry as a moderator and mediator of emotion", *Biological Psychology*, vol. 67, no. 1-2, 2004, pp.7-49.

[8] Y. Liu, O. Sourina, and M.K. Nguyen, "Real-time EEG-based emotion recognition and its applications", *Transactions on computational science XII*, vol. 6670, 2011, pp. 256-277.

[9] J.A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, 1980, pp.1161-1178.

[10] T. Zikov, S. Bibian, G. A. Dumont, M. Huzmezan, and C. Ries, "A wavelet based de-noising technique for ocular artifact correction of the electroencephalogram", in: *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2002, pp. 98-105.

[11] R.J. Jennrich, "Stepwise discriminant analysis", in: *Statistical Methods for Digital Computers*, edited by K. Einslein, K. Ralston, and H.S. Wilf New York: Wiley, 1977.

[12] L.I. Aftanas and S.A. Golocheikine, "Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution EEG investigation of meditation," *Neuroscience Letters*, vol. 310, no. 1, 2001, pp. 57-60.

[13] D.J. Schutter and J. van Honk, "Decoupling of midfrontal delta-beta oscillations after testosterone administration," *International Journal of Psychophysiology*, vol. 53, no. 1, 2004, pp. 71–73.

[14] D. Sammler, M. Grigutsch, T. Fritz and S. Koelsch, "Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music," *Psychophysiology*, vol. 44, no. 2, 2007, pp. 293-304.

[15] P. Putman, J. van Peer, I. Maimari, and S. van der Werff, "EEG theta/beta ratio in relation to fear-modulated response-inhibition, attentional control, and affective traits," *Biological Psychology*, vol. 83, no. 2, 2010, pp. 73-78.

[16] L. Sebastiani, A. Simoni, A. Gemignani, B. Ghelarducci, and E.L. Santarcangelo, "Autonomic and EEG correlates of emotional imagery in subjects with different hypnotic susceptibility," *Brain Research Bulletin*, vol. 60, no. 1-2, 2003, pp. 151-160.

[17] L.I. Aftanas, N.V. Reva, L.N. Savotina, and V.P. Makhnev, "Neurophysiological correlates of induced discrete emotions in humans: an individually oriented analysis," *Neuroscience and Behavioral Physiology*, vol. 36, no. 2, 2006, pp. 119-130.

[18] W.J. Ray and H.W. Cole, "EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes", *Science*, vol. 228, no. 4700, 1985, pp. 750-752.

[19] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Research Reviews*, vol. 29, no. 2-3, 1999, pp. 169-195.

[20] L.I. Aftanas and S.A. Golocheikine, "Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution EEG investigation of meditation," *Neuroscience Letters*, vol. 310, no. 1, 2001, pp. 57-60.

[21] B.L. Fredrickson and D. Kahneman, "Duration neglect in retrospective evaluations of affective episodes," *Journal of Personality and Social Psychology*, vol. 65, no. 1, 1993, pp. 45–55.

[22] D. Kahneman, "Evaluation by moments, past and future," in: D. Kahneman and A. Tversky (Eds.) *Choices, Values and Frames*, Cambridge University Press, 2000, pp. 693-708.