1F5-OS-06b-1

EEG-Based Emotion Recognition during Music Listening

Nattapong Thammasan[*1] Ken-ichi Fukui[*2] Koichi Moriyama[*2] Masayuki Numao[*2]

[*1] Graduate School of Information Science and Technology, Osaka University
[*2] The Institute of Scientific and Industrial Research (ISIR), Osaka University

In Music Emotion Research, the goal is to quantify and explain how music influences our emotional states. Scientists realize that human brain is relevant to emotion, and it leads to the study of human emotion by capturing information from brain. Electroencephalogram (EEG) is an efficient tool to capture brainwave. This research proposes a framework to recognize human emotions during music listening by using EEG. In this research, MIDI music files are used, 12 electrodes of EEG are selected and placed according to the 10-20 international standard. Two-Dimensional emotion model is used to represent human emotions. The experiment starts from music selection, music listening with brainwave capturing, and ends with emotion self-annotation continuously. For data analysis, Fractal Dimension value calculations by Higuchi Algorithm are performed, and Support Vector Machine is applied. The results of emotion classification are satisfied, with the accuracy of 90% for arousal classification and 86% for valence classification on average.

Keywords: Electroencephalogram, Fractal Dimension, Support Vector Machine

1. Introduction

It is known that music can induce human emotion. In Music Emotion Research, the goal is to quantify and explain how music influences our emotional states. Human emotion is quite difficult to study and related to many things such as physiology, biological reactions, or mental states. Recently, scientists realized that human brain was relevant to human emotion, and it leaded to studying of human emotion and feeling by using information from brain, including brainwave. Therefore, by studying brainwave, we could get some cue about current emotion of human. Then, it is possible to determine human emotion during music listening, which could be an important link to understand or explain the influence mechanism of music to human state.

Meanwhile, electroencephalogram, a tool to capture brainwave, is rather efficient tool, and becomes more reasonable in cost. Emotion recognition during listening to music, therefore, by using electroencephalogram (EEG) becomes active in recent years. However, it is still in early state research with many challenges. The lack of standards of data collection methodology and sensitivity of EEG signal to noise of disturbing environment are the issue that researchers need to deal with. Moreover, the emotion tags from another experiment were used in previous research, where they could be not related to true emotion of subjects.

This research is to propose automatic emotion recognition while subjects are listening to music by using EEG. MIDI file type is selected to use as it allows high-level analysis of music feature. Subjects' self-reports are utilized for model training and testing, which they could reflect true emotion of subject during music listening.

The next challenge is to discover a feature that can be utilized to improve accuracy of classification. Familiarity level to individual song is one of interesting candidate based on the hypothesis that familiar song can induce person more emotionally because of background of that song.

2. Related Works

Human emotion can be described in concrete way both quantitatively and qualitatively. Plutchik proposed categorized emotion model, in which the eight basic emotions are defined as follows: anger, fear, sadness, disgust, surprise, anticipation, acceptance and joy [Plutchik 2003]. After that, love, awe, disappointment, and another emotions are added to be basic emotions. In 1979, Russell proposed dimensional continuous emotion model, or bipolar model of emotion [Russell 1979]. Here, two dimensions such as arousal (strong level) and valence (positive/negative emotion) were used to describe emotions. Then, controllable ability was added to be third dimension of emotion model as well. Recently, the dimensional emotion classification is more widely used in the emotion recognition research because emotions such as joy, surprise, etc. can be located in the dimensional space. However, dimensional emotion model such as arousal-valence space model is still used in emotion recognition research in present day due to its simple and effectiveness in emotion representation.

Researchers proposed automatic emotion recognition systems by using EEG. In [Bos 2006], sound clips from International Affective Digitized Sounds (IADS) and images from International Affective Picture System (IAPS) are used and emotion states can be classified into positive/arousal, positive/calm, negative/arousal, and negative calm with accuracy of 92.3%. In [Lin 2010] work, pre-labeled emotional music are used, and joy, sadness, anger, pleasure emotions can be differed with the performance rate 85%. In [Sourina&Liu 2011], researcher used both self-report emotion after listening to sound clips and emotion labeled from clips of IADS and positive/ high-aroused, positive/low-aroused, negative/high-aroused, and negative/low-aroused emotions can be discriminated with the accuracy of 84.9% for arousal classification and 90% for

Contact: Nattapong Thammasan, Department of Architecture for Intelligence, The Institute of Scientific and Industrial Research (ISIR), 8-1, Mihogaoka, Ibaraki, Osaka 567-0047, Japan, E-mail: Nattapong@ai.sanken.osaka-u.ac.jp

valence classification. However, previous researches were lacked of effective emotion annotation from participant, which could be more directly related to brain signal than the emotion tagged by another experiments.

In our approach, we realize the importance of subject's emotion annotation, which is done continuously in each song clip. Two-dimensional emotion model is selected to represent human emotion.

3. Research Methodology

3.1 Data Collection

The experiment starts with questionnaire about personal information and music related questions for subjects. Then, subject selects MIDI files to listen in experiment. The songs in playlist are composed of familiar 3 songs and 3 unfamiliar songs. 12 Utilized electrodes are placed around frontal lobe of the brain according to the International 10–20 Standard. After that, subject listens to all full-length-songs, meanwhile brainwave is captured by Waveguard™EEG [Waveguard]. 15 seconds between songs are included for relaxing and avoiding effects from previous song. At the end, subject listens to previous songs again, and annotates current emotion simultaneously by clicking on a location in two-dimensional emotion space shown on monitor screen continuously. Subject can comment about the experiment after all sessions are finished.

3.2 Data Analysis

The main processes of creating emotion recognition model is composed of feature extraction from EEG signal and classification based on targeted emotions. There are various algorithms proposed to recognize emotion from EEG signals to improve performance of feature extraction and classification. In feature extraction step, Fractal Dimension values of EEG calculated by Higuchi algorithm [Higuchi 1988] is selected because of its simplicity and satisfied performance rate of overall basic emotion classification in two-dimensional space in previous work [Sourina&Liu 2011].

The Higuchi algorithm calculates fractal dimension value from time-series data. For example, $X(1)$, $X(2)$, ..., $X(N)$ is a finite set of time series. We can construct time series as following definition:

$$X_k^m = X(m), X(m+k), ..., X\left(m + \left[\frac{N-m}{k}\right] \cdot k\right), \qquad (1)$$

(m = 1, 2, ..., k) where m is the initial time and k is the interval time, then k sets of $L_m(k)$ are calculated as follows:

$$L_m(k) = \frac{1}{k} \cdot \left\{ \left(\sum_{i=1}^{\left[\frac{N-m}{k}\right]} \left| X(m+ik) - X(mi(i-1) \cdot k) \right| \right) \frac{N-1}{\left[\frac{N-m}{k}\right] \cdot k} \right\}, \qquad (2)$$

where $\langle L_m(k) \rangle$ denotes the average value of $L_m(k)$, and a relationship exists as follows:

$$\langle L_m(k) \rangle \propto k^{-D} \qquad (3)$$

Then, the fractal dimension can be obtained by logarithms plotting between different k and its associated $\langle L_m(k) \rangle$.

Asymmetry of signal power between left and right hemisphere would be used for determining valence value due to corresponding with positive/negative emotion. Exited state level of a subject could be obtained by using information gained from electrodes and can be used for arousal classification [Sourina&Liu 2011]. For emotion classification, non-linear classifier will be used due to nature of brainwave signal data. Support vector machine is used in this step.

We applied Higuchi Algorithm to raw signal data obtained from every electrodes. We use asymmetry of FD values from Fp1-Fp2, F3-F4, C3-C4, Fz-Pz, F7-F8, T3-T4 electrode pairs as features for valence classification, and we use FD values calculated from Fp1, Fp2, F3, F4, C3, C4, Fz, Pz, F7, F8, T3, T4 electrodes as features for arousal classification. Moreover, emotion annotated by subjects are divided to 4 classes, which are positive/high-aroused, positive/low-aroused, negative/high-aroused, and negative/low-aroused emotions accordings to position in arousal-valence models. The emotion class labels are added to data for classification by support vector machine algorithm.

4. Results

Data from 3 subjects with 20669 instances of data is inputted for classifying and testing with 10-fold cross validation testing mode. Each data element has duration of 0.004 seconds. The sliding windows size is 1000, with sliding step of 50 (0.2 seconds). We can classify arousal level with the correctness of 86-95%, and we can classify valence level with the correctness of 74-99% (Fig. 1).

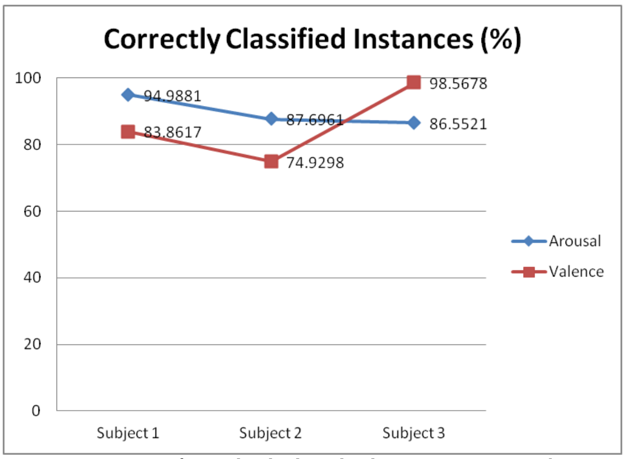

Figure 1: Comparison of arousal and valence levels recognition accuracy between subject

After analyzing the effect of familiarity level (Fig. 2), we can find that using familiar songs merely can improve accuracy of valence classification in all subjects, and unfamiliar songs usage can improve accuracy of arousal classification in all subjects.

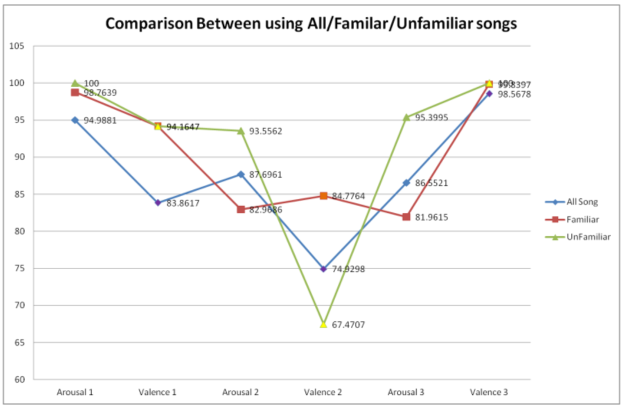

Figure 2: The effect of using familiar songs and unfamiliar songs merely on arousal and valence levels recognition accuracy

5. Discussion

In this experiment, the number of subjects is quite limited. However, the separated classification of arousal and valence class can give us high performance. More experiments is needed for further investigation.

The result is also comparable to similar work of [Yamano 2012], where the researchers can estimate arousal and valence values of emotion with 84.4% and 87.3% of accuracy of arousal classification and valence classification respectively.

The effect of familiarity is not significant and convincing enough for any conclusion because of few number of participants. However, in emotion recognition model creation, using both familiar songs and unfamiliar songs could be considered for further experiment setting up.

6. Conclusion

In this research, we propose a framework to create automatic emotion recognition model by using brainwave captured from Electroencephalogram (EEG). The MIDI songs are used as emotional stimuli, which are composed of 3 familiar songs and 3 unfamiliar songs in each subject. 12 Electrodes are placed according to International 10-20 Standard. Higuchi algorithm is applied to calculate fractal dimension value from EEG data. Continuously emotion self-annotated emotion by subjects is used as labels for classification. From the data of 3 subjects, arousal and valence classification are performed separately with high performance. We can classify the arousal value with 86-95% accuracy and the valence value with 74-99% accuracy. Furthermore, familiar songs can improve accuracy of valence classification and unfamiliar songs usage can improve accuracy of arousal classification in all subjects. More experiments are needed where the factor of familiarity in song selection in future experiments should be considered.

References

[Plutchik 2003] Plutchik, R.: Emotions and life : perspectives from psychology, biology, and evolution. American Psychological Association, Washington, DC, 1st edition, 2003.

[Russell 1979] Russell, J.A. Affective Space is bipolar. Journal of Personality and Social Psychology, 37, 345-356, 1979.

[Bos 2006] Bos, D.O. EEG-based emotion recognition. The influence of Visual and Auditory Stimuli[Online], 2006. Available: http://hmi.ewi.utwente.nl/verslagen/capita-selecta/CS-Oude_Bos-Danny.pdf

[Lin 2010] Yuan-Pin Lin; Chi-Hong Wang; Tzyy-Ping Jung; Tien-Lin Wu; Shyh-Kang Jeng; Jeng-Ren Duann; Jyh-Horng Chen, "EEG-Based Emotion Recognition in Music Listening," Biomedical Engineering, IEEE Transactions on , vol.57, no.7, pp.1798,1806, July 2010 doi: 10.1109/TBME.2010.2048568

[Sourina&Liu 2011] Olga Sourina, Yisi Liu, A Fractal-based Algorithm of Emotion Recognition from EEG using Arousal-valence model. In Proc. Biosignals 2011, Rome, 26-29 Jan, pp.209-214, 2011.

[Higuchi 1988] Higuchi, T. Approach to an irregular time series on the basis of the fractal theory, Physica D: Nonlinear Phenomena, 31, 277-283, 1988.

[Yamano 2012] Yu Yamano, Rafael Cabredo, Paul Salvador Inventado, Roberto Legaspi, Koichi Moriyama, Ken-ichi Fukui, Satoshi Kurihara, and Masayuki Numao. "Estimating Emotions on Music Based on Brainwave Analyses", Proc. The 3rd International Workshop on Empathic Computing (IWEC-2012) (held in conjunction with PRICAI 2012), Sarawak, Malaysia, Sep. 3-4 2012.

[Waveguard] http://www.ant-neuro.com/products/waveguard