Emotion Recognition Based on Low-Cost In-Ear EEG

Gang Li*, Zhe Zhang, Guoxing Wang, *Senior Member*, *IEEE*

***Abstract*—In this paper, we propose a low-cost in-ear EEG device which is implemented by refitting a commercial scalp EEG device, in order to recognize emotion in a manner that is simple, inexpensive, and popular in style. EEG signals of twelve subjects were recorded under three emotion conditions that were induced by music and video materials. By using wavelet packet transformation (WPT), two frequency features and a nonlinear feature are extracted to create a three-dimensional feature vector for each labeled EEG segment. These feature vectors are input into a support vector machine (SVM) classifier for automatic emotion recognition. The SVM classifier achieved a best 94.1% cross-validation accuracy for positive (high valence, HV) and negative (low valence, LV) two-class emotion recognition. However, the accuracy for excited (high valence and high arousal, HVHA), relaxed (high valence and low arousal, HVLA) and negative (LV) multi-class emotion classification was 58.8%. The experimental results show that the proposed low-cost in-ear EEG has outstanding accuracy for valence recognition, but poor accuracy for arousal recognition.**

***Index Terms*— In-Ear EEG, Emotion Recognition, WPT, SVM**

I. INTRODUCTION

UNDERSTANDING human emotions is a key factor in the computer's ability to understand human behavior and to communicate intelligently with people, many human-machine interaction (HMI) studies focused on human emotion recognition, in order to make HMI more efficient and more natural.

Over the past three decades, researchers have already started from different behavior modalities (such as face [1], voice [2], posture [3], or a combination of these modalities [4,5]) to study emotion recognition. However, using behavior methods to recognize emotion has a major drawback when the emotion remains intentionally invisible (ie, no any facial expressions and gestures). Therefore, in recent years more and more researchers began to explore the relationship between physiological signals and emotions [6-8]. Physiological signals used in these studies were recorded from autonomic nervous system in the periphery, such as ECG, skin resistance, EMG, etc. In addition to these periphery physiological signals, signals captured from the brain in central nervous system have been proved to provide informative characteristics in responses to the emotional states. For example, EEG signals are produced by brain neuronal electrical activity, which are typically acquired non-invasively via electrodes placed on the scalp. Compared to other physiological signals, EEG signals are more directly related to emotion-related nerve changes. Thierry Pun *et al.* analyzed the correlates of EEG and peripheral physiological signals on the assessment of emotions induced by video games. Experimental results show that EEG has a better accuracy than skin resistance, photoplethysmograph, respiration and body temperature [9]. Actually, EEG has been widely used in cognitive neuroscience to investigate the regulation and processing of emotion for the past decades [10-12]. However, these studies were based on cumbersome and expensive EEG devices, which poses a major practical hurdle when detecting emotion in daily-life style.

The concept of in-ear EEG was proposed in 2012 by Mandic *et al.* [13] in Imperial College of Technology. Crucial advantages of in-ear EEG compared to existing scalp EEG are summarized as follows (adapted from [13]):

- The in-ear EEG electrode is discreet and easy to put in place by the users themselves, facilitating everyday use.
- The tight fit between the in-ear EEG electrode and the ear canal guarantees that the electrodes are held firmly in place, diminishing motion artifacts.
- Muscle artifacts are greatly reduced as there are no muscle fibers in the ear canal and common sources of muscle artifacts, such as the face and eye muscles, are located far away.
- The ear canal is a cavity enclosed by an electrical conductive medium (skin, fluids, brain tissue), reducing the interference from external electrical fields. Interference from external magnetic sources is also diminished as the area spanned by the measurement loop of the leads is very small.

Mandic *et al.* suggested that the design of the in-ear EEG sensor is consistent with the traditional scalp EEG sensor, that is, the same electrode material, amplifier and acquisition mode (ie, the same reference and ground electrode position). In [13], the authors used sintered Ag/AgCl as the in-ear EEG electrode material. The reference electrode and the ground electrode were placed in the right earlobe and the chin respectively. During the experiment, the in-ear EEG electrode needs to be embedded on a customized hearing aid, and then apply the conductive gel to the electrode. In 2016, K. S. Park *et al.* from Seoul National University and J. Chuang *et al.* from University of California shared their latest research findings on the in-ear EEG applications at the annual conference on IEEE Engineering in Medicine and Biology Society [14, 15].

Manuscript received on June 25, 2017. This research is sponsored by Shanghai Sailing Program under Grant 17YF1426900.

G. Li, Z. Zhang and G. Wang are with the Department of Micro/Nano Electronic Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China (*corresponding author: G. Li; e-mail: rivermandy@sjtu.edu,cn).

978-1-5090-5803-7/17/$31.00 ©2017 IEEE

K. S. Park *et al*. conducted an experiment to detect driver drowsiness using in-ear EEG. During the experiment, they adopted the metal electrode and conductive gel to sense the in-ear EEG, and used a research-class EEG device (BIOPAC Systems, Inc) to complete the signal amplification, filtering, analog-to-digital conversion and wireless transmission. J. Chuang *et al.* tested the performance of mental tasks using in-ear EEG. They used gold cup electrode with conductive gel to sense in-ear EEG, and a low-cost EEG device (Neurosky, Inc) to complete the follow-up signal processing. Table I summarizes all the existing brain machine interface (BMI) applications on in-ear EEG. We can see clearly that human's emotion recognition using the in-ear EEG has not been tested yet.

TABLE I
THE EXISTING RESEARCHES ON IN-EAR EEG AND SCALP EEG

Application field			By Scalp EEG	By In-Ear EEG
BMI	Active	Mental Task	√	[15]
		Motor Imagery	√	-
	Reactive	SSVEP(1)	√	[13]
		ASSR(2)	√	[13]
	Passive	Fatigue Monitoring	√	[14]
		Emotion Recognition	√	-

(1) Steady state visually evoked potential (2) Auditory steady state response

The objective of this paper is to uncover the association between the in-ear EEG dynamics and emotions by 1) designing emotion-inducible experiments, 2) searching emotion-specific features of the in-ear EEG and 3) testing the accuracy of different SVM classifiers. Ultimately, this study attempts to lay the foundation for a wearable, inexpensive and easy-to-use emotion detector.

II. METHODS

A schematic of the proposed approach has been shown in Fig. 1. The process of in-ear EEG based emotion recognition is divided into three stages: -Stage I: data collection, Stage II: feature extraction and Stage III: feature classification. The Stage I involves the setting up of in-ear EEG device and the completion of emotion-inducible experiment. The Stage II includes WPT-based signal decomposition and feature extraction. The Stage III involves the feature classification using SVM classifier with different kernels. The following sections introduce in detail the three stages respectively.

A. Data Collection

In-ear EEG data in this study were collected from 12 healthy undergraduate and graduate students (10 males and 2 females; age 22-26) during music listening and video watching. The selection of music and video materials is according to DEAP database experiment [16], where excited emotion stimulated by an encouraging video about athletes and rock music, relaxed emotion stimulated by scenery video and *What a wonderful world*, and negative emotion by car accident video and terrible noisy music *Lost rivers*. The in-ear EEG device, as shown in Fig. 2, was designed and implemented by refitting Neurosky Mindset EEG headset. Its sampling rate is 512Hz and contains a 3-100Hz bandpass filter. It is important to note that before finally settling with Neurosky, other low-cost EEG

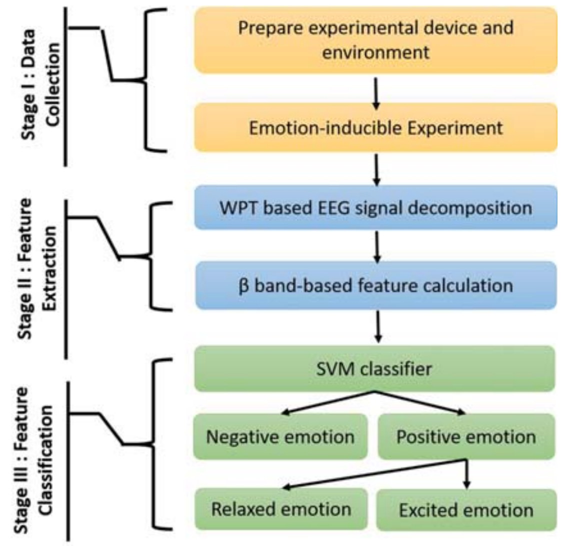

Fig. 1. Flowchart of in-ear EEG-based emotion recognition.

devices, such as Emotiv (Emotiv Inc) and Muse (InteraXon Inc), were compared in the aspects of retail price and equipped software. Neurosky Mindset is not only the most inexpensive one, but also supports OpenVibe which is an open source software for real-time BMI application and raw data acquisition in this study. Similar to [15], the modifications made to the original Neurosky Mindset included releasing the electrode from the plastic forehead arm, removing the electrode, and replacing it by soldering a new 10mm disc-shaped Ag/AgCL electrode onto the original wire. An elastic earplug was used to support the electrode and make it tight fit to the upper edge of the ear canal.

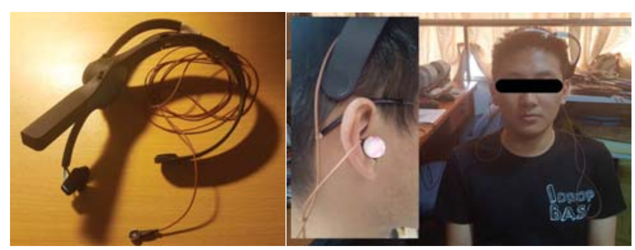
Fig. 2. The proposed low-cost in-ear EEG device and its real application for one of the subjects.

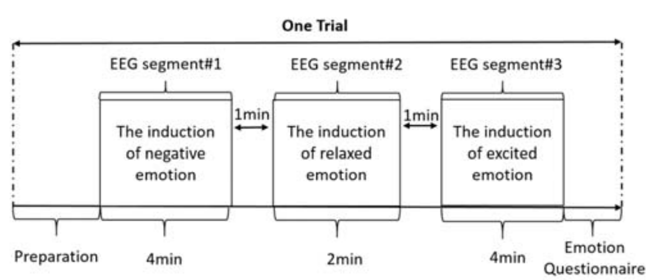

Fig. 3. The data collection procedure

The data collection procedure, as shown in Fig.3, began with a preparation session including an informed consent process and a set up period with the in-ear EEG device, followed by a set of three emotion-inducible materials presented on a desktop while in-ear EEG was recorded, and finally a post-

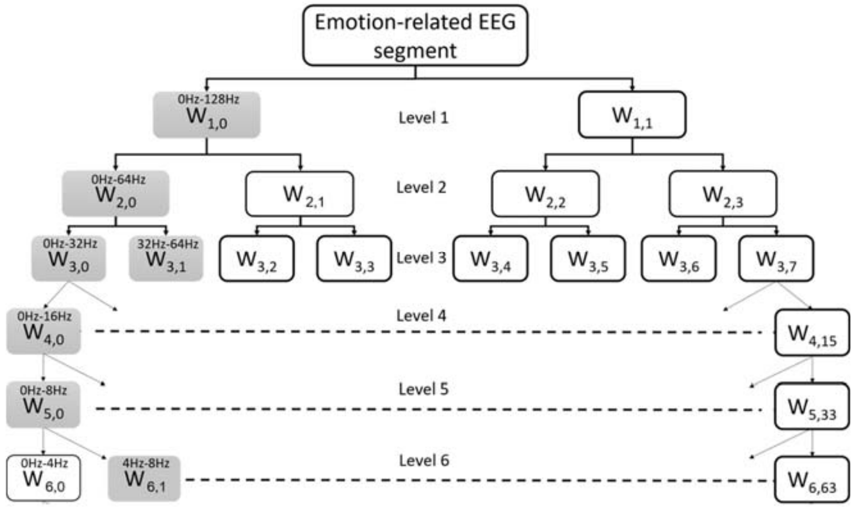

Fig. 4. Illustration of 6-level wavelet packet transform and corresponding frequency components. The white boxes components are filtered through setting the coefficients to zero. Boxes in the middle of level 4, 5 and 6 are omitted.

experiment questionnaire. What we expected was that the three emotion-inducible materials could make subjects feel negative emotion (low valence, LV), relaxed (high valence and low arousal, HVLA) and excited (high valence and high arousal, HVHA), according to J. Russel's *circumplex model* [17]. If the self-reported emotion was different from our expected induced emotion or reported as 'not strong during the experiment', the subject would be tested again in another day. If the induced emotions of the subject were still different with our expected in the second testing, the in-ear EEG data of the subject would not be regarded as valid samples.

B. Feature Extraction

Assuming that the EEG signal is non-stationary, wavelet transform can be applied to extract the EEG features. The regular wavelet decomposition method, such as discrete wavelet transform (DWT), decomposes the given signal into a set of approximate and detailed coefficients for the n^{th} level (shown as shaded blocks in Fig. 4). Each block (the wavelet coefficient) can be defined as a frequency component $W_{n,m}$, which corresponds to the range of $[\frac{mf_s}{2^{n+1}}Hz, \frac{(m+1)f_s}{2^{n+1}}Hz]$, where f_s is the sampling rate, which is 512Hz in this study and m=0,1,…2^n-1. WPT decomposes not only the approximate coefficients, but also the detailed coefficients. Therefore, the information (high frequency) which is lost in DWT can be retrieved by using WPT. In this study, WPT decomposes each emotion-related EEG segments into eight levels, which is based on the coiflets mother wavelet with order 5. On the basis of the frequency component $W_{n,m}$, the widely-used EEG frequency bands are obtained (e.g., frequency components $W_{8,14}$-$W_{8,29}$ corresponds to the β band) and then the relative power spectrum energy of β band (denoted by P_β), the variance of β band (denoted by Var_β), and the sample entropy of β band (denoted by $SamEn_\beta$) are calculated. Finally, the normalized Var_β, P_β and $SamEn_\beta$ are fed into the SVM classifier, where the normalized Var_β is calculated by $Var_{\beta\_relaxed}/Var_{\beta\_negative}$ and $Var_{\beta\_excited}/Var_{\beta\_negative}$.

C. Feature Classification

The SVM is used to classify the in-ear EEG features. The primary advantage of SVM is its ability to minimize both structural and empirical risk, thereby leading to better generalizations for new data classifications, even with limited training datasets. In this study, a binary SVM is trained to classify negative and positive emotion (including relaxed and excited emotions) first, and then a three-class SVM is trained to classify negative, relaxed and excited emotions respectively.

$$K(x,y) = \exp\left(\frac{-\|x-y\|^2}{\gamma}\right) \quad (1)$$

The LibSVM [18] implementation for SVM was used in this study. The constant parameter C and γ were chosen through grid-search and one with the best cross-validation accuracy (C =100,000, γ =1) were picked. The radial basis function (RBF) kernel given by Equation 1 was selected for classifying the features after comparing with linear and polynomial kernel. Leave-one-subject-out cross-validation [19] was carried out to calculate the classification accuracy.

III. Results and Conclusion

Totally, thirty-four EEG segments out of the thirty-six segments (12 subjects x 3 emotion-related segments) were valid samples. One self-reported bad emotion segment and one bad signal segment is excluded. Therefore, thirty-four feature sets comprising Var_β, $SamEn_\beta$ and P_β were extracted. As can be seen in the Box-Whiskers plots (Fig. 5), the normalized Var_β shows a linear increasing trend with the increase of valence, however that is not the case for $SamEn_\beta$ and P_β. This was the motivation behind trying a non-linear classifier such as SVM.

The extracted features are then fed into the RBF-based SVM classifier. The results of RBF kernel with different parameters C and γ are shown in Table II. As it can be seen, the accuracy ranges from 67.6% to 97.1% for negative-positive binary classification and 47.1% to 58.8% for negative-relaxed-

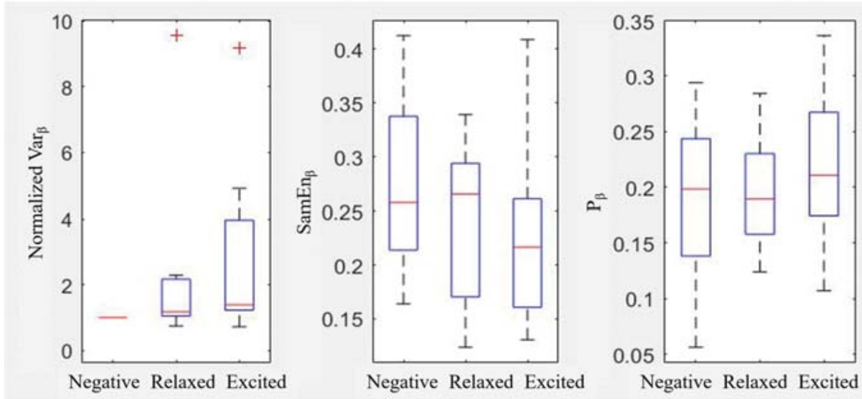

Fig. 5. Box-Whiskers plots of the in-ear EEG features that are extracted from the three emotion-related EEG segments. The boxes have three lines to present the values for first quartile (the bottom), median, and third quartile (the top) for column data. The length between the first quartile (Q1) and the third quartile (Q3) is called interquartile range (IQR). Two addition lines at both ends of the whisker indicate the Q1 − 1.5 × IQR and Q3 + 1.5 × IQR value of a column data. Any data not included between the whiskers is plotted as outliers represented by "+". The number next to the outlier is the number of the data in that column, called case number

excited multi-class classification. The best accuracy was achieved by using C=10^5 and γ =1. In comparison, the accuracy of binary classification is better than multi-class classification. A further classification test between excited and relaxed state shew 59% accuracy, verifying the decrease in accuracy for arousal recognition. The self-reported relatively low level of stimulated excited emotion in data collection experiment and the own mechanism of in-ear EEG are two proposed reasons for the decrease. The experiment indicates the feasibility of single-channel in-ear EEG for emotion recognition, but under the limitation of poor arousal recognition. Further research will use more sophisticated device with better experiments that can fully arouse the emotion to detect the feasibility of in-ear EEG emotion recognition. Self-adaptive feature extracting methods such as sparse auto encoder is also a substitution for artificial feature extracting with expected better result.

TABLE II
EMOTION RECOGNITION ACCURACY USING IN-EAR EEG, WPT-BASED FEATURES AND SVM

	γ=1 (default)			
Classification Type	C=10	C=10^3	C=10^4	C=10^6
Binary	67.6%	85.3%	94.1%	94.1%
Multi-class	47.1%	55.9%	58.8%	55.9%

References

[1] A. J. Fridlund, P. Ekman, and H. Oster, "Facial Expressions of Emotion: Review Literature 1970-1983," *Nonverbal Behavior and Communication*, A.W. Siegman and S. Feldstein, eds., pp. 143-224. Hillsdale NJ: Lawrence Erlbaum Assoc., 1987

[2] T. Johanstone, R. Banse, and K.S. Scherer, "Acoustic Profiles in Prototypical Vocal Expressions of Emotions," *Proc. Int'l Conf. Phonetic Science*, vol. 4, pp. 2-5, 1995.

[3] R.J. Holt, T.S. Huang, A.N. Netravali, and R.J. Qian, "Determining Articulated Motion from Perspective Views," *Pattern Recognition*, vol. 30, no. 9, pp. 1435-1,449, 1997

[4] L.S. Chen, T.S. Huang, T. Miyasato, and R. Nakatsu, ªMultimodal Human Emotion/Expression Recognition," Proc. Int'l Conf. Automatic Face and Gesture Recognition, pp. 366-371, 1998.

[5] G. Zen, L. Porzi, E. Sangineto, E. Ricci, N. Sebe, "Learning Personalized Models for Facial Expression Analysis and Gesture Recognition,"IEEE Trans on Multimedia, Vol. 18, No. 4, pp.775-788, April 2016

[6] J. Kim and E. Andre, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.

[7] K. H. Kim, S. W. Bang, and S. R. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Med. Biol. Eng. Comput.*, vol. 42, no. 3, pp. 419–427, May 2004.

[8] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 10, pp. 1175–1191, Oct. 2001.

[9] G. Chanel, C. Rebetez, M. Betrancourt and T. Pun, "Emotion assessment from physiological signals for adaption of game difficulty, " *IEEE Trans on systmes, man, and cybernetics-Part A: Systems and humans*, vol. 41, No. 6, pp. 1052-1063, Nov.2011.

[10] Y. P. Lin *et al*., "EEG-based emotion recognition in music listening, " *IEEE Trans on Biomedical Engineering*, vol. 57, no. 7, pp. 1798-1806, July 2010

[11] W. L. Zheng, B. L. Lu, "Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks," IEEE Trans on Autonomous Mental Development, Vol.7, No.3, pp.162-175,May 2015.

[12] J. H. Zhang, M. Chen, S. Q. Hu, Y. Cao, R. Kozma,"PNN for EEG-based Emotion Recognition,"Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, pp. 2319-2323, Oct 9-12 2016, Budapest, Hungary, Feb 2017.

[13] D. Looney, P. Kidmose, C. Park, M. Ungstrup, M.L.Rank, K.Rosenkranz, D.P.Mandic,"The In-the-Ear Recording Concept: User-Centered and Wearable Brain Monitoring,"IEEE Pluse, pp.32-42,Vol.3, No.6, Nov 2012.

[14] T. H Hwang, M.Y Kim, S.Y.Hong, K.S.Park, "Driver drowsiness detection using the in-ear EEG," Proceedings of IEEE 38th Annual International Conference on the Engineering in Medicine and Biology Society, Orlando, USA, 16-20 Aug,2016.

[15] M. T. Curran, J. K. Yang, N. Merrill, J. Chuang, "Passthoughts Authentication with Low Cost EarEEG,"Proceedings of IEEE 38th Annual International Conference on the Engineering in Medicine and Biology Society, Orlando, USA, 16-20 Aug,2016.

[16] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, ": A Database for Emotion Analysis Using Physiological Signals," IEEE Trans on Affective Computing, pp.18-31, Vol.3, No.1, Jan 2012.

[17] J. A. Russell *et al*. "Relating the Five-Factor Model of Personality to a Circumplex Model of Affec," in *The Five-Factor Model of Personality Across Cultures*. Springer Press, US, 2002:79-104.

[18] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf (accessed on 25 June 2017).

[19] G. Li and W. Y. Chung, "A context-aware EEG headset system for early detection of driver drowsiness," *Sensors,* vol. 15, no. 8, pp. 20873-20893, August 2015.