

Auditory evoked potentials in patients with major depressive disorder measured by Emotiv system

Dongcui Wang^{a,b}, Fongming Mo^a, Yangde Zhang^a, Chao Yang^c, Jun Liu^d, Zhencheng Chen^{b,*} and Jinfeng Zhao^{a,*}

^aXiangya Hospital, Central South University, No. 87 Xiangya Road, Changsha 410008, China

^bSchool of Life and Environmental Science, Guilin University of Electronic Technology, No. 1 Jinji Road, Guilin 541004, China

^cSchool of Electronic Engineering and Automation, Guilin University of Electronic Technology, No. 1 Jinji Road, Guilin 541004, China

^dThe Social Welfare Hospital of Guilin, No. 89 North Zhongshan Road, Guilin 541001, China

Abstract. In a previous study (unpublished), Emotiv headset was validated for capturing event-related potentials (ERPs) from normal subjects. In the present follow-up study, the signal quality of Emotiv headset was tested by the accuracy rate of discriminating Major Depressive Disorder (MDD) patients from the normal subjects. ERPs of 22 MDD patients and 15 normal subjects were induced by an auditory oddball task and the amplitude of N1, N2 and P3 of ERP components were specifically analyzed. The features of ERPs were statistically investigated. It is found that Emotiv headset is capable of discriminating the abnormal N1, N2 and P3 components in MDD patients. Relief-F algorithm was applied to all features for feature selection. The selected features were then input to a linear discriminant analysis (LDA) classifier with leave-one-out cross-validation to characterize the ERP features of MDD. 127 possible combinations out of the selected 7 ERP features were classified using LDA. The best classification accuracy was achieved to be 89.66%. These results suggest that MDD patients are identifiable from normal subjects by ERPs measured by Emotiv headset.

Keywords: Emotiv system, major depressive disorder, event-related potential

1. Introduction

Depression, characterized by having deficits in motivation, concentration and executive functions, is one of the most common mental disorders. Traditionally, the clinical diagnosis of depression is mainly based on subjective symptomatology. With the emergence of novel technologies detecting weak bio-signals, measurable data could be generated as supporting evidence for diagnosis of Major Depressive

* Address for correspondence: Zhencheng Chen, School of Life and Environmental Science, Guilin University of Electronic Technology, No. 1 Jinji Road, Guilin, China. Tel.: +86 15295885965; Fax: +86 0773-2292232; E-mail:chenzhcheng@163.com.

Jinfeng Zhao, Xiangya Hospital, Central South University, No. 87 Xiangya Road, Changsha, China. Tel.: +86 15874856567; Fax: +86 0731-84327987; E-mail: zhaojinfeng@hotmail.com.

Disorder (MDD).

Event-related potentials (ERPs) are a time locked measure of brain electrical activity. Due to its objectivity, this measure is well suited to study brain activities and has been applied to subjects suffered from depression for decades. ERPs is an additive signal from several overlapping components and subcomponents like N400, Contingent Negative Variation (CNV), mismatch negativity(MMN), N1, N2 and P3. Many studies have focused on amplitudes, latencies and topographic distributions of these components [1, 2], especially the N1, N2 and P3 in depressed patients. It is well known that the characteristics of ERPs represent a distinct pattern of cognitive mechanisms that rely on pre-attentive, executive and affective processes. To be specific, N1 subcomponent, occurring at around 100 ms post-stimulus, is involved in perception or initial sensory processing and also linked to early selective attention; the N2, occurring at nearly 200 ms post-stimulus, is affected by executive mechanisms; and the P3, occurring in the range of 270-550 ms, is regarded as an index of allocation of processing resources, orientation of attention to salient or novel stimuli, memory recall and response modulation.

The amplitudes of N1 and N2 have been proposed as important features associated with depression. However, there were controversial findings in previous studies. Several studies have reported N1 reductions in depressed subjects [3], but more studies have failed to find distinctive N1 differences between the depressed patients and the normal controls [4, 5]. Some studies found increased N2 amplitude [6], and others reported no difference [7, 8] or even reduced N2 in depressed subjects [9]. P3 measure, which is used to investigate cognitive deficits in depressed patients, has achieved much consistent findings, including reduced P300 amplitude and abnormal P300 latency [9].

As to Emotiv system, although comparative studies have reported its poorer fidelity compared with laboratory EEG recording systems like ANT [10] and Neuroscan [11], it still could provide valuable information on EEG data. And supporting arguments have been given by a number of researchers who have successfully applied this system to assess cognitive processes [12-17].

The current study aimed to explore the potential utilization of the Emotiv headset on clinical MDD patients. To our knowledge, this would be the first report of how Emotiv headset performs on patients.

2. Methods

2.1. Subjects

Participants were all right-handed, literate and without auditory and/or visual impairments. Normal subjects were recruited from Guilin University of Electronic Technology. Only those who met the following criteria were designated as the control group: 1) scored less than 10 on Beck Depression Inventory (BDI, 13 items); 2) no personal or familial history of psychiatric disorders, including claustrophobia; 4) no history of drug or substance abuse; 5) no history of taking any tranquilizers or antidepressants; and/or 6) no consumption of coffee or alcohol in the last 24 hours. In sum, fifteen participants were selected to be the control (7 men and 8 women, aged from 23 to 27 years (25.3 ± 1.1)).

Sixty-six outpatients from Guilin Social Welfare Hospital who received consistent diagnoses of depression from two professional Psychiatrists based on Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) participated in data acquisition. However, twelve of them were in comorbid with schizophrenia, thirty recrudescence patients had not gone through drug washout period when data acquisition was performed, one was left-handed, and one's behavioral response data were missing due to an operational error, leaving only 22 qualified for the depression group. To match with the control, 15 out of 22 were selected to represent the depression group (7 men and 8 women, aged from 17 to 42

years (30.1±8.0)).

This study was approved by the regional ethics committee of Guilin University of Electronic Technology and Guilin Social Welfare Hospital. Written approvals were collected from all participants. Data were anonymized before data analysis.

2.2. Task and procedure

Monotonous stimuli were designed using GoldWave software (www.goldwave.ca/). They were 50ms-long sine waves with frequencies of 500 Hz (standard stimulus, number of appearances = 190) and 1,000 Hz (standard stimulus, number of appearances = 30). These tones were broadcasted binaurally in a pseudo-random order. Adjacent appearances of target stimulus were avoided. The inter-stimulus interval was randomly generated within a range of 800ms to 1,200 ms. To be well prepared, the subject was instructed to adjust himself to his most comfortable position and put his fingertips on the basic positions for touch typing. For performing the experiment, the subject was told to make responses only to the target stimulus by pressing the SPACE button and silently counting the total appearances of the target stimulus. The procedure was initiated with a practicing session. The subject was permitted to enter the experimental session only when his response accuracy exceeded 85%. Every trial started with a visual fixation cross followed by a stimulus tone and ended with a full black screen. The experimental session lasted for about six minutes.

2.3. Data recording

A research edition of wireless Emotiv EPOC headset (www.emotiv.com) was used. Electrophysiological data were acquired from 10 channels: F7, F3, T7, P7, O1, O2, P8, T8, F4, and F8. These sensors were placed according to the International 10/20 placement. Electrical resistance was reduced by saline liquid. Data were online referenced at CMS/DRL, sampled at 128 Hz, bandpass filtered from 0.2 to 45 Hz, and notched filtered at 50 Hz and 60 Hz.

Behavioral information was recorded synchronously with brain signal data, from which response accuracy and RT measures were derived. Specifically, response accuracy was determined by the ratio of the number of correct responses over the number of total target stimuli; RT referred to the time interval between the target stimulating event and the corresponding response event.

2.4. Data processing

Data were processed under MATLAB environment with the help of functions from EEGLAB toolbox (<http://sccn.ucsd.edu/eeglab/>) [18]. EOG artifact was reduced using a combined method of extended informax-ICA (without PCA) and Multiple Artifact Rejection Algorithm (MARA) [19]. Data were then epoched from -200 ms to 1000 ms and baseline was corrected by referring to pre-stimulus interval. For epoch rejection, an amplitude threshold of ±75 uv was configured. Across epochs survived the amplitude screening, and with qualified key-press responses, data were averaged.

For each participant and each component of interest, N1, N2 and P3, individual peak amplitudes and corresponding latencies across the following three intervals: 70 – 170 ms, 180 – 270 ms and 280 – 550 ms, were extracted from the averaged ERP.

Analysis of variance (ANOVA) was taken for statistical analyses with alpha value of $P < 0.05$ as the significant difference level.

Table 1
Behavioral results of Emotiv control and Emotiv depressed groups

Characteristic	Emotiv Control (N = 15)	Emotiv Depressed (N = 15)	Mean Difference (Depressed - Control)
RT ^a (ms)	549.7 ± 72.5	620.8 ± 123.7	71.1*
Response accuracy (%)	99.1 ± 1.54	95.1 ± 10.1	-4.0

Note: ^a Reaction Time; * indicates the significant difference ($P < 0.05$) by ANOVA; Values were mean ± SD.

2.5. Feature selection and classification

Liu and Motoda [20] introduced the algorithm of Relief-F for feature selection by which, initiative instance was randomly chosen and made comparisons with its nearest neighbor, after repetitive calculations and comparisons, the weights of the relevance among instances were conducted. This method has become one of the most useful strategies in feature selection. In the present study, the features consisted of RTs, response accuracy, amplitudes and latencies of N1, N2 and P3 at all ten electrodes, 62 pieces in total. To eliminate the effect of randomness on initial instance, Relief-F was run for 20 times and each time on resembled half instances from each group, and the number of the most important features was estimated according to dataset redundancy criteria. Linear Discriminant Analysis (LDA) is a common classification technique. Using all the possible combinations of the selected features, the classification accuracy of LDA with leave-one-out cross-validation was calculated.

3. Results and discussion

The control and the depressed group were configured with similar gender ratio. But ages were significantly different. Thus, the possible effects of gender and age on both behavioral and ERP measures were also carefully examined by repeat ANOVA tests.

3.1. Behavioral performance

ANOVAs performed on RT between the control and the depressed group (Table 1) showed a significant difference ($F(1,28) = 4.67$, $p = 0.03$). Depressed patients responded to stimuli much slower than the normal subjects. Response accuracy data were lower in the depressed group in general, but there was no statistical difference by ANOVA test ($F(1,28) = 2.31$, $p = 0.14$). Neither gender nor age exerted any effect on RT and response accuracy in the control or depressed group.

3.2. Latency and amplitude of ERP component N1, N2 and P3

In N1, the depressed group had shorter latencies than the control group. Repeat ANOVAs revealed a significant difference at the electrode F3 ($F(1,28) = 7.00$, $P = 0.01$), F4 ($F(1,28) = 5.35$, $P = 0.02$) and O1 ($F(1,28) = 4.78$, $P = 0.03$), and a tendency of difference at F8 ($F(1,28) = 4.15$, $P = 0.05$), but not at the other six electrodes. Regarding amplitudes of N1, the depressed group generally had lower N1 amplitudes than the control. Significant difference was shown at F4 ($F(1,28) = 5.82$, $P = 0.02$), O1 ($F(1,28) = 4.90$, $P = 0.03$) and O2 ($F(1,28) = 4.72$, $P = 0.03$).

In N2, the latencies of depressed group were not different from the control group by one way ANOVA. The absolute N2 amplitudes in depressed group were generally higher. However,

statistically significant differences were only seen at electrodes F3 ($F(1,28) = 4.88$, $P = 0.03$) and F4 ($F(1,28) = 5.80$, $P = 0.02$).

In P3, the depressed group took longer average latencies than the control group at all electrodes, but statistically significant differences were only presented at F7 ($F(1,28) = 6.59$, $P = 0.01$), T7 ($F(1,28) = 7.41$, $P = 0.01$), P8 ($F(1,28) = 4.27$, $P = 0.04$), O1 ($F(1,28) = 7.02$, $P = 0.01$), and O2 ($F(1,28) = 7.30$, $P = 0.01$). Likewise, the depressed group showed increased amplitudes at all electrodes, especially at the electrodes F4 ($F(1,28) = 4.36$, $P = 0.04$), T8 ($F(1,28) = 9.77$, $P = 0.00$), P7 ($F(1,28) = 10.31$, $P = 0.00$), P8 ($F(1,28) = 7.04$, $P = 0.01$), O1 ($F(1,28) = 10.66$, $P = 0.00$), and O2 ($F(1,28) = 9.66$, $P = 0.00$).

Gender and age had no significant effects on latencies and amplitudes of N1, N2 or P3.

3.3. Selection and classification of typical features of MDD

Based on the principle of least redundancies, seven features were selected from the 62 features sorted by Relief-F and shown as follows: 1) amplitude of N1 at electrode F4, 2) amplitude of N2 at F4, 3) amplitude of P3 at O1, 4) latency of N1 at F3, 5) latency of N2 at T7, 6) latency of P3 at T7, and 7) RT. All the possible combinations based on the above 7 features were tested for classification accuracies using LDA classifier with leave-one-out cross-validation method. And the results ranged from the highest classification accuracy of 89.66% with 1 combination, 86.21% with 14 combinations, and 82.76% with 9 combinations to the lowest 10.34%.

3.4. Discussion

The present study was characterized by the abnormal features of ERPs in depressed patients that may be applicable for clinical diagnosis as measurable evidence. It is also shown that Emotiv headset was able to discriminate the abnormal EPRs of MDD patients from the normal subjects. Thus there would be commercial value to exploit Emotiv headset as a household medical device.

The Emotiv EPOC headset had been scrutinized on normal subjects for its capability in detecting electrophysiological signals [11, 21]. In the present study, the Emotiv headset was applied to depressed subjects. Behavioral data indicated that patients with depression took more time to response.

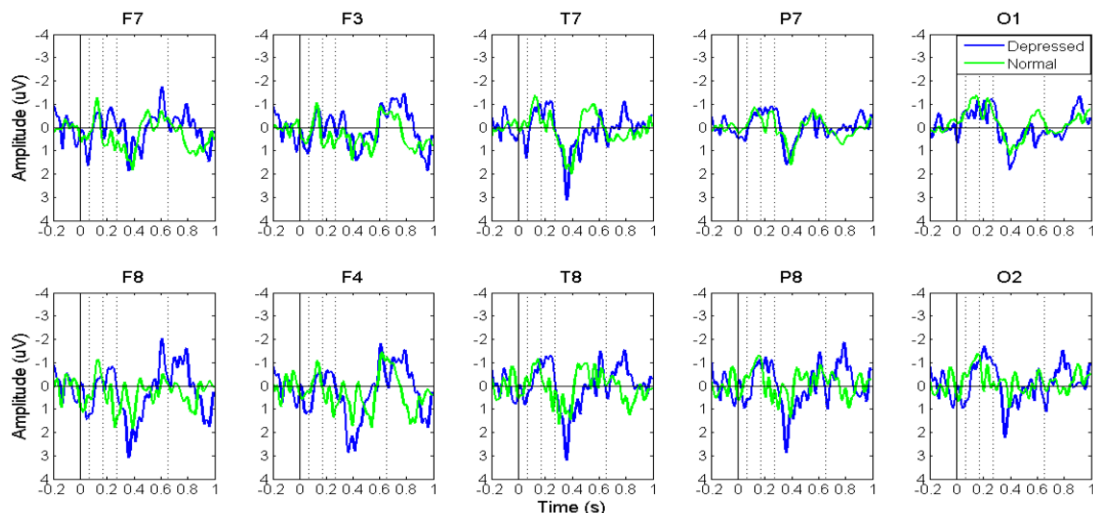


Fig. 1. Typical grand average ERPs of normal (in green) and depressed group (in blue) measured by Emotiv system.

These findings were consistent with the depression symptoms that typically showed patients' difficulty in concentration and their indecisiveness. Therefore, the RT was considered as an important behavioral component for defining the character of depression. In analysis of electrophysiological signals, prolonged latencies were found in depressed group in ERP components N2 and P3 (Figure 1). Our findings on latencies were supported by other studies. Urretavizcaya, et al. [9] declared that latencies in N1, N2 and P3 were all significantly longer in depressed patients than those in the normal subjects. Reduced amplitudes of N1 and increased amplitudes of N2 and P3 were also revealed in depressed group (Figure 1). It had been reported that the impairments in selective attention was in association with weakened N1 in depressed patients [3]. Sur and Sinha [22] also mentioned a negative relationship between N2 amplitude and persistence score, implicating higher N2 amplitude versus a lower persistence score. Our results based on data from Emotiv system reached a consensus on these two findings; however they failed to replicate observations from Urretavizcaya, et al. [9] who had declared an association between decreased P3 amplitude and major depression.

Among many classification methods, such as support vector machine, k -nearest neighbors, naïve Bayesian algorithms [23, 24], LDA was chosen for the merits of its simple application and of having been well characterized in other discriminant studies [25]. LDA in current study explicitly attempted to model the difference between the normal and depressive subjects with the percentage of classification accuracy to indicate the degree of difference. The classification results showed that the combination of (1, 3) respectively corresponding to the amplitude of N1 at F4 and amplitude of P3 at O1 was a very important factor because it not only achieved the highest classification accuracy (89.66%), but also was presented in 9 out of 15 combinations counted by over 85% classification accuracy. The combination of (2, 3) also occurred frequently with 8 occurrences in 15 combination models. Interestingly, feature 1 and feature 2 both happened at frontal F4, which was reported strongly associated with depression [26]. The site of O1 locates in occipital hindbrain. It is very interesting that 64 combinations above 68.97% comprised feature 3, indicating that the site of O1 may reflect certain electrophysiological changes relating to depression.

To sum up, the LDA study suggests some significance and implications: 1. The ERPs in MDD patients display some typical differences from the normal subjects and 7 features are highlighted. 2. The normal and depressed ERPs captured by Emotiv headset can be analyzed and discriminated using LDA method. 3. LDA illustrates 15 models of feature combinations with high classification accuracies in which the abnormal ERP features in MDD patients can be identified by a more precise and accurate method. Hopefully, the LDA study on data collected by Emotiv headset has implications for developing some MDD diagnostic products.

Acknowledgments

This study was supported by Guangxi Experiment Center of Information Science at Guilin University of Electronic Technology (Grant No. 20130106), Guangxi Key Lab of Automatic Detection Technology and Instrument (Grant No. YQ14116) and National Science and Technology Support Program (Grant No. 2013BAI03B01). The authors are grateful to Hong-Bo Chen, Chun-Hua Wang, Shao-Feng Mo, Shao-Hui Luo, Yang-Guang Deng, Rong-Jing Li, Yan-Ning Huang, Tian Gu, Lin-Ling Zhou and Qing-Hong Ye.

References

- [1] S. Sur and V.K. Sinha, Event-related potential: An overview, *Industrial Psychiatry Journal* **1** (2009), 70–73.
- [2] N.C. Zygouris, Event related potentials in depressed children and adolescents, prolonged P300 Latency and abnormal brain asymmetry, *Neuroscience Letters* **500** (2011), e20–e21.
- [3] M.A. Burkhart and D.G. Thomas, Event-Related potential measures of attention in moderately depressed subjects, *Electroencephalography and Clinical Neurophysiology* **1** (1993), 42–50.
- [4] T. Linka, G. Sartory, S. Bender, et al., The intensity dependence of auditory ERP components in unmedicated patients with major depression and healthy controls, an analysis of group differences, *Journal of Affective Disorders* **103** (2007), 139–145.
- [5] C.E. Tenke, J. Kayser, S.A. Shankman, et al., Hemispatial PCA dissociates temporal from parietal ERP generator patterns: CSD components in healthy adults and depressed patients during a dichotic oddball task, *International Journal of Psychophysiology* **1** (2008), 1–16.
- [6] G.E. Bruder, C.E. Tenke, J.P. Towey, et al., Brain ERPs of depressed patients to complex tones in an oddball task: Relation of reduced P3 asymmetry to physical anhedonia, *Psychophysiology* **1** (1998), 54–63.
- [7] D.H. Blackwood, L.J. Whalley, J.E. Christie, et al., Changes in auditory P3 event-related potential in schizophrenia and depression, *British Journal of Psychiatry* **150** (1987), 154–160.
- [8] S. Kaiser, J. Unger, M. Kiefer, et al., Executive control deficit in depression: Event-related potentials in a Go/Nogo task, *Psychiatry Research* **3** (2003), 169–184.
- [9] M. Urretavizcaya, I. Moreno, L. Benlloch, et al., Auditory event-related potentials in 50 melancholic patients: Increased N100, N200 and P300 latencies and diminished P300 amplitude, *Journal of Affective Disorders* **3** (2003), 293–297.
- [10] M. Duvinage, T. Castermans, M. Petieau, et al., Performance of the Emotiv EPOC headset for P300-based applications, *Biomedical Engineering Online* **12** (2013), 56.
- [11] N.A. Badcock, P. Mousikou, Y. Mahajan, et al., Validation of the Emotiv EPOC® EEG gaming system for measuring research quality auditory ERPs, *PeerJ* **1** (2013), e38.
- [12] N. Jatupaiboon, S. Pan-ngum and P. Israsena, Emotion classification using minimal EEG channels and frequency bands, 2013 10th International Joint Conference on Computer Science and Software Engineering (Jesse), Maha Sarakham, 2013, pp. 21–24.
- [13] N. Jatupaiboon, S. Pan-ngum and P. Israsena, Real-time EEG-based happiness detection system, *Scientific World Journal* **2013** (2013).
- [14] T.D. Pham and D. Tran, Emotion recognition using the emotiv EPOC device, *Neural Information Processing 2012* **7667** (2012), 394–399.
- [15] E.W. Anderson, K.C. Potter, L.E. Matzen, et al., A user study of visualization effectiveness using EEG and cognitive load, *Computer Graphics Forum* **3** (2011), 791–800.
- [16] E.T. Esfahani and V. Sundararajan, Classification of primitive shapes using brain-computer interfaces, *Computer-Aided Design* **10** (2012), 1011–1019.
- [17] S. Fok, R. Schwartz, M. Wronkiewicz, et al., An eeg-based brain computer interface for rehabilitation and restoration of hand control following stroke using ipsilateral cortical physiology, 2011 Annual International Conference of the Ieee Engineering in Medicine and Biology Society (EMBC), Boston, MA, 2011, pp. 6277–6280.
- [18] A. Delorme and S. Makeig, Eeglab: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, *Journal of Neuroscience Methods* **1** (2004), 9–21.
- [19] I. Winkler, S. Haufe and M. Tangermann, Automatic classification of artifactual ICA-components for artifact removal in EEG signals, *Behavioral and Brain Functions* **1** (2011), 30.
- [20] H. Liu and H. Motoda, Computational methods of feature selection, in: *Data Mining and Knowledge Discovery Series*, Chapman & Hall/CRC, Boca Raton, 2007, pp. 440.
- [21] C. Kranczioch, C. Zich, I. Schierholz, et al., Mobile EEG and its potential to promote the theory and application of imagery-based motor rehabilitation, *International Journal of Psychophysiology* **1** (2014), 10–15.
- [22] S. Sur and V.K. Sinha, Auditory P300 study in remittent bipolar child and adolescent population, *Asian Journal of Psychiatry* **4** (2009), 128–131.
- [23] W. Bosl, A. Tierney, H. Tager-Flusberg, et al., EEG complexity as a biomarker for autism spectrum disorder risk, *BMC Medicine* **1** (2011), 1–16.
- [24] I. Habes, S.C. Krall, S.J. Johnston, et al., Pattern classification of valence in depression, *Neuroimage Clinical* **2** (2013), 675–683.
- [25] K. Okamoto and Y. Harasawa, Prediction of symptomatic depression by discriminant analysis in japanese community-dwelling elderly, *Archives of Gerontology and Geriatrics* **2** (2011), 177–180.
- [26] A.M. Mitchell and P. Pospel, Frontal brain activity pattern predicts depression in adolescent boys, *Biological psychology* **2** (2012), 525–527.