35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Multi-domain Feature Analysis for Depression: a Study of N170 in Time, Frequency and Spatial Domains Simultaneously

Qiangfeng Zhao, Junfeng Sun, *Member, IEEE*, Fengyu Cong, *Member, IEEE*, Shan Chen, Yingying Tang, and Shanbao Tong, *Senior Member, IEEE*

***Abstract*—This study investigated the differences in event-related potentials (ERPs) between depression and normal control groups by using the cue-target paradigm with facial expressions as stimuli. Conventional ERPs analysis did not show a significant difference in the N170 amplitude or latency between the two groups. However, the multi-domain feature analysis of N170 by nonnegative tensor factorization (NTF) indicated that N170 in depression group had lower power compared with the normal control group for all three different emotional (i.e., happy, neutral, and sad) facial stimuli ($p \leq 0.05$). The results revealed the perceptual early-stage dysfunction in face processing for depression.**

I. Introduction

Depression is a mental disorder that is characterized by low mood, low self-respect and losing interests in usually enjoyable activities. Facial expression of human is responsive to social interactions, and this responsiveness may account for effects of social support on depression [1]. Thus, facial expressions are widely used as stimuli in experiments on cognition of mental disorder, particularly the depressed patients. Behavioral studies have indicated that depressed patients exhibit blunted processing of emotional expressions in general [2]. Gur and colleagues found that depressed patients had a poorer behavior across tasks, and were less sensitive in the task of happy and sad emotion discrimination [2]. Nevertheless, the results on processing of emotional expressions in depression so far are not so consistent. A recent study supported that depression was more sensitive to negative facial expression [3].

The N170, a well-known event-related potential (ERP) component in EEG signals, is face-specific and could be elicited by human faces [4]. One study reported that the N170 was unaffected by any emotional expression, supporting the hypothesis that structural encoding and expression analysis are independent processes [5]. However, another study reported a different result that an early automatic encoding of emotional expressions was reflected by N170 [6]. The role of N170 in emotional facial expression cognition is in debate.

ERPs have been widely used in the study of depression. However, most studies only analyzed ERPs based on their peak measurements (e.g., amplitude and latency) in time domain. Such information in one domain is inevitably affected by many factors, such as the variance within participant groups (e.g., age, race and gender), resulting in heterogeneous estimations of ERP components. Such a problem may be overcome by multi-way signal processing methods by reducing effects of exogenous factors for ERP analysis.

Recently, one of the multi-way signal processing methods, nonnegative tensor factorization (NTF), conforming to the canonical polyatic (CP) model [also coined as Parallel Factor Analysis (PARAFAC) model] [7], has been used to extract the multi-domain feature of an ERP by exploiting its information in the time, frequency and spatial domains simultaneously [8, 9]. The multi-domain feature analysis of ERP shows strength in discriminating different groups of participants compared with the conventional peak amplitude and latency of ERP [8, 9]. In this study, hence, based on the CP model, we use NTF to extract the multi-domain features of N170 for discriminating depression and normal groups. It should be noted that this study is entirely different from [8, 9] in purposes of the conducted research, experimental designs, participants, and data collection and preprocessing. Here, we just apply the data processing method in [8, 9] where the method was proposed.

II. Materials and Methods

A. Subjects and Stimuli

Twenty-two depressed patients (denoted as DEP hereafter, age: 31.7 ± 9.5 years; male/female: 12/10), recruited from Shanghai Mental Health Center, and twenty-two age-matched normal control subjects (denoted as NOR, age: 33.9 ± 8.7 years; male/female: 9/13), recruited from local community, participated in this study. Informed consent was obtained after complete description of the study and all subjects were compensated for their participation regardless their performance. All experiments were performed in a sound-attenuating and electrically shielded room. Totally, four kinds of pictures, i.e., 12 happy, 12 neutral and 12 sad facial pictures selected from Nimstim facial picture set [10], and 5 control pictures of objects (e.g., desk), were used as stimuli. All pictures were converted into gray color and the hair in each facial picture was cut off (see the example in Fig. 1) [11]. The presentation of stimuli was controlled by E-prime 2.0.

*This work was supported by National Natural Science Foundation of China (No. 60901025), Med-X Research Fund of SJTU (No. YG2011MS68), and SJTU Joint Science and Engineering Funding.

Q. Zhao (e-mail: zhaoqf1@sjtu.edu.cn), J. Sun (e-mail: jfsun@sjtu.edu.cn), S. Chen (e-mail: chenshan@sjtu.org), and S. Tong (corresponding author; phone: +86-21-34205138; fax: +86-21-34204717; e-mail: stong@sjtu.edu.cn) are with the School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, 200240, P. R. China.

F. Cong (e-mail: fengyu.cong@jyu.fi) is with the Department of Mathematical Information Technology, University of Jyväskylä, Finland.

Y. Tang (e-mail: yytang0522@yahoo.com.cn) is with Department of EEG Source Imaging, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, P. R. China.

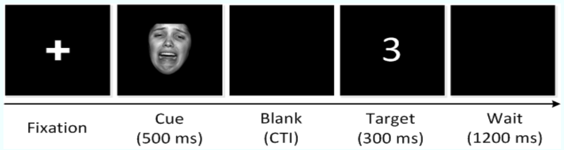

Figure 1. The cue-target paradigm [11].

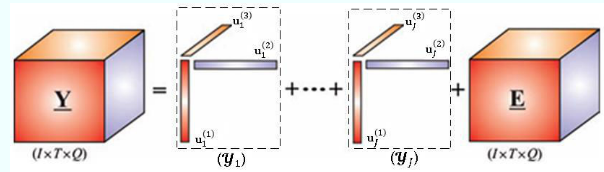

Figure 2. A graphical representation of the third-order tensor $\underline{\mathbf{Y}}$ as a sum of J rank-one tensors $\boldsymbol{\mathcal{Y}}_j$ under CP model [7].

B. Experimental Procedures

The cue-target paradigm is shown in Fig. 1. In each trial, a cue, which may be one of the happy, the sad, the neutral facial pictures, or the control object pictures, was first presented at the center of screen for 500 ms. Then a black blank was shown for a certain period, which is called cue-target interval (CTI). Following that, a digit target (3 or 5) was presented for 300 ms with equal appearance probabilities. Participants were asked to pay attention to the emotion of the facial cue, and then discriminate the target by pressing the left mouse button for digit 3 and the right mouse button for digit 5 as soon as possible. However, there was no need to make response in the case of object cues, and this was to ensure the participants' concentration. Three different CTIs (350 ms, 1000 ms, and 1500 ms) were used in different trials with equal appearance probabilities. The test contained 6 blocks with 120 trials each.

During the experiment, continuous EEG signals were recorded using an elastic cap with 32 scalp electrodes (EasyCap, Brain Products, Germany) at 1000 Hz sampling rate. Scalp impedance at each electrode was kept below 10 kΩ, and all EEG signals were referenced to the tip of nose channel and on-line digitally filtered into 0.5 to 100 Hz.

C. Data preprocessing

We used Brain Vision Analyzer 2.0 (Brain Products, Germany) to process the EEG data offline. The vertical and horizontal EOGs were used to detect eye movements and blinks. After artifact removal, 60 valid EEG trials of the same CTI and the same type of cues were extracted and further averaged to get the corresponding ERPs. We defined the ERP from 250 ms before the onset of cue-stimulus to 1000 ms after the offset of cue-stimulus as "cue part". The baseline was corrected based on the average amplitude of the 250 ms pre-stimulus period. In this paper, we only focused on the "cue part" of ERP of the case with CTI = 1500 ms, and the influences of different CTIs on emotion recognition will be presented in a separated study.

To extract the multi-domain feature by NTF, the ERP signals of all subjects were firstly transformed by Morlet wavelet into the fourth-order time-frequency tensor [8], denoted as $\underline{\mathbf{Y}}$ hereafter. The dimensions of $\underline{\mathbf{Y}}$ were frequency (74 bins from 0.4 to 15 Hz) by time (1750 samples from -250 to 1500 ms) by space (30 channels) by subject (44).

D. Introduction of NTF

NTF conforming to the CP model is illustrated in Fig. 2 [7] for a third-order tensor. According to NTF [7], a N^{th}-order $\underline{\mathbf{Y}}$ can be expressed as the sum of J rank-one tensors $\boldsymbol{\mathcal{Y}}_j$ as:

$$\underline{\mathbf{Y}} = \underline{\widehat{\mathbf{Y}}} + \underline{\mathbf{E}} = \sum_{j=1}^{J} \boldsymbol{\mathcal{Y}}_j + \underline{\mathbf{E}}, \quad (1)$$

where $\underline{\mathbf{E}}$ is the error of the approximation; $\boldsymbol{\mathcal{Y}}_j$ $(j = 1,2,\cdots,J)$ are the output rank-one tensors of NTF. Each $\boldsymbol{\mathcal{Y}}_j$ can be further expressed as the outer product of vectors: $\boldsymbol{\mathcal{Y}}_j = \mathbf{u}_j^{(1)} \circ \mathbf{u}_j^{(2)} \circ \cdots \circ \mathbf{u}_j^{(N)}$, where N is the number of modes of $\underline{\mathbf{Y}}$. $N = 3$ is the case in Fig. 2 and $N = 4$ in this study. Most algorithms for NTF are to minimize a squared Euclidean distance as the following cost function [7]

$$D\left(\underline{\mathbf{Y}}|\underline{\widehat{\mathbf{Y}}}\right) = \frac{1}{2}\left\|\underline{\mathbf{Y}} - \underline{\mathbf{I}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \cdots \times_N \mathbf{U}^{(N)}\right\|_F^2. \quad (2)$$

$\mathbf{U}^{(n)} = \left[\mathbf{u}_1^{(n)}, \mathbf{u}_2^{(n)}, \cdots, \mathbf{u}_J^{(n)}\right]$ $(n = 1,2,\cdots,N)$, for the $\mathrm{n^{th}}$ mode, is the component matrix. Details for approximating the $\underline{\widehat{\mathbf{Y}}}$ can be found in [7].

E. NTF for ERPs

In this study, the fourth-order tensor $\underline{\mathbf{Y}}$ for ERP signals contains four factors, i.e., frequency, time, space and subject. Therefore, Eq. (1) becomes [8]

$$\underline{\mathbf{Y}} \approx \underline{\widehat{\mathbf{Y}}} = \sum_{j=1}^{J} \mathbf{u}_j^{(f)} \circ \mathbf{u}_j^{(t)} \circ \mathbf{u}_j^{(c)} \circ \mathbf{f}_j. \quad (3)$$

where, $\mathbf{f}_j$ denotes the subject factor and is named as the j^{th} multi-domain feature [8] extracted by NTF, and each participant has one value. $\mathbf{u}_j^{(f)}$, $\mathbf{u}_j^{(t)}$, and $\mathbf{u}_j^{(c)}$ are spectral, temporal and spatial components corresponding to the frequency, time and space factors. According to the CP model, $\mathbf{u}_j^{(f)}$, $\mathbf{u}_j^{(t)}$, and $\mathbf{u}_j^{(c)}$ reveal the spectral, temporal and spatial properties of the multi-domain feature $\mathbf{f}_j$.

For NTF, we applied the hierarchical alternating least squares algorithm [7] and nonnegative constrains are added for each factor during the decomposition. Note that this algorithm has high stability of decomposed results under random initialization in the application of multi-domain feature extraction for ERPs [8].

The number of output rank-one tensors $\boldsymbol{\mathcal{Y}}_j$, i.e., J, is an important parameter for NTF in extracting the multi-domain feature of an ERP because the selection of different J could lead to different results [8]. Here, we analyzed the NTF models with J from 1 to 60. Then, DIFFIT method [12], based on criteria of variance of $\underline{\mathbf{Y}}$, was applied to estimate the proper number of rank-one tensors for the NTF algorithm.

After J multi-domain features of ERPs are extracted, the desired one(s) can be determined according to prior knowledge and statistical tests [8]. Since N170 is of interest here, therefore, the multi-domain features with the temporal components peaking around 170 ms were selected. Then, statistical tests were performed on each selected features to find the desired one(s) with significant difference between two groups.

F. Statistical Analysis

To examine the difference between depression and normal groups, repeated-measures ANOVA was used for the conventional ERPs analysis. For the extracted multi-domain features by NTF, the statistical analysis was implemented on a 22 by 2 feature data matrix with $p < 0.05$ as significance. We used one-way ANOVA and Kruskal-Wallis test, following the manner in [8]. This is because multi-domain features by NTF might not meet the assumptions of ANOVA. Kruskal-Wallis test can overcome the drawback of ANOVA to some extent [13]. The minimal p-value of the two tests was reported to show the group difference in extracted multi-domain features.

III. RESULTS

A. Conventional ERPs Analysis

A three-factor (electrode × cue × group) repeated-measures ANOVA were performed on N170 amplitude and latency, which both showed a main effect of cue (i.e., happy, neutral, sad facial pictures, and pictures of control objects) ($p < 0.001$). Further test showed the N170 amplitude of the cases of happy, neutral, sad faces was larger than that of control object pictures ($p < 0.001$) and the N170 latency of the cases of control object pictures was larger than those of three different faces ($p < 0.001$). Nevertheless, within-group test did not show significant difference among three facial stimuli ($p > 0.1$). Furthermore, the ANOVA showed main effect of group was under a significance level ($p > 0.1$), although depressed subjects tended to have overall lower N170 amplitude than normal control group (DEP: -6.365 ± 5.736 μV; NOR: -7.860 ± 5.578 μV).

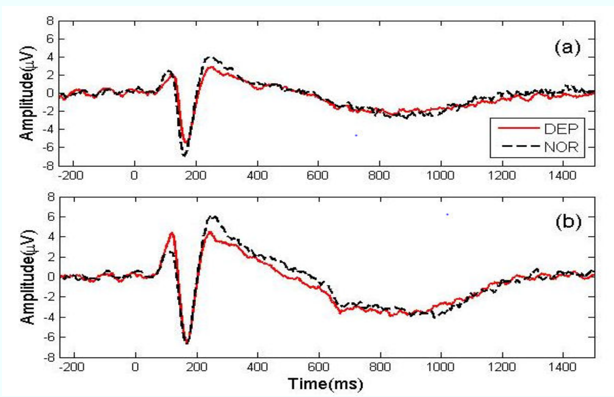

Figure 3. Grand averaged ERPs for depression (DEP) and normal control (NOR) groups for the case of CTI = 1500 ms and happy facial cues at sites P7 (a) and P8 (b).

B. NTF analysis of ERP

To investigate the effect of emotional facial stimuli to depression, the NTF method was performed only on the "cue part" of ERPs of the three facial cues except the ERPs for the cues of control object pictures.

In the case of happy face cue, the number of rank-one tensors estimated by DIFFIT [12] was 36, and one desired multi-domain feature showing significant difference between two groups [F(1,43) = 8.915, p = 0.0047] among 36 ones extracted by NTF is shown in Fig. 4(a), as well as the corresponding temporal, spectral and spatial components in Figs. 4(b)-4(d). Obviously, the temporal peak latency is around 170 ms in the temporal component [Fig. 4(b)] and the spectrum peak is around 5 Hz in the spectral component [Fig. 4(c)]. Furthermore, the spatial component [Fig. 4(d)] shows an occipito-temporal distribution and has maximal amplitude at the lateral posterior channels (e.g., P8/P7). All these match the properties of N170 [14]. Thus, we attribute this multi-domain feature to N170. According to Fig. 4(a), this desired multi-domain feature indicates that the depression group had significantly lower power of N170 compared with the normal control group. In the case of neutral face cue, the number of rank-one tensors determined by DIFFIT was 41. The desired multi-domain feature of N170 [F(1,43) = 4.073, p = 0.0500] and its corresponding components are presented in Figs. 4(e)-4(h). In the case of sad face cue, the number of rank-one tensors was also 41 by DIFFIT. The desired multi-domain feature of N170 [F(1,43) = 5.717, p = 0.0214] and its corresponding components are presented in Figs. 4(i)-4(l). We can observe that the selected desired multi-domain features and their corresponding components for the cases of happy, neutral, and sad face cues were similar. The desired multi-domain feature of N170 in depression group had significantly lower power compared with normal control group in all cases.

IV. DISCUSSION

In this study, conventional ERPs analysis showed face-sensitive N170 compared with the non-face objects confirming the result in [4], which also suggested that different emotion of facial expressions elicited different responses in brain and N170 was affected as well [4]. Stekelenburg and Gelder reported that expressions of anger and fear were associated with an enhanced N170 [15], and Krombholz *et al.* also reported similar results [16]. However, some other studies reported that N170 was unaffected by emotional facial expressions and suggested that spatial attention could modulate the structural encoding of faces [17]. In this study, we also found that the amplitude of N170 had no significant diversity between the cases of emotional face and the case of neutral face, and there was no bias specific to sad faces. In addition, conventional ERPs analysis also showed that there was no significant difference between depression group and normal control group on N170 amplitude and latency in all cases. Nonetheless, we found lower power of multi-domain feature of N170 in depression by NTF.

One of the major challenges in studying N170 was to represent N170 correctly regarding the complexity of participants' EEG. The N170 amplitude and latency used in conventional ERP analysis are based on the information of the ERP waveform observed only in the time domain. However, with information of brain responses in multiple domains (i.e., the time, frequency, and spatial domains) used simultaneously, the multi-domain feature of N170 extracted by NTF is expected to be more reliable. As a result, the difference in N170 between the depression group and the normal control group could be identified by NTF but failed to be detected by conventional ERPs analysis.

As N170 is functionally referred to the structural analysis of face components, the lower power of multi-domain feature of N170 extracted by NTF may reflect a perceptual early-stage dysfunction in face processing in depressed patients. Similar result, that is, a reduced N170 amplitude in response to faces, had been reported in schizophrenia patients [18]. On the contrary, a recent research reported that N170 had a higher peak amplitude at P7 in depressed patients while responding to the oddball stimuli (with low appearance probability) [9]. The inconsistent results may be due to the different experiment paradigms, as our study used the equalball stimuli. Moreover, here, depressed patients had lower power of multi-domain feature of N170 in response to all types of faces,

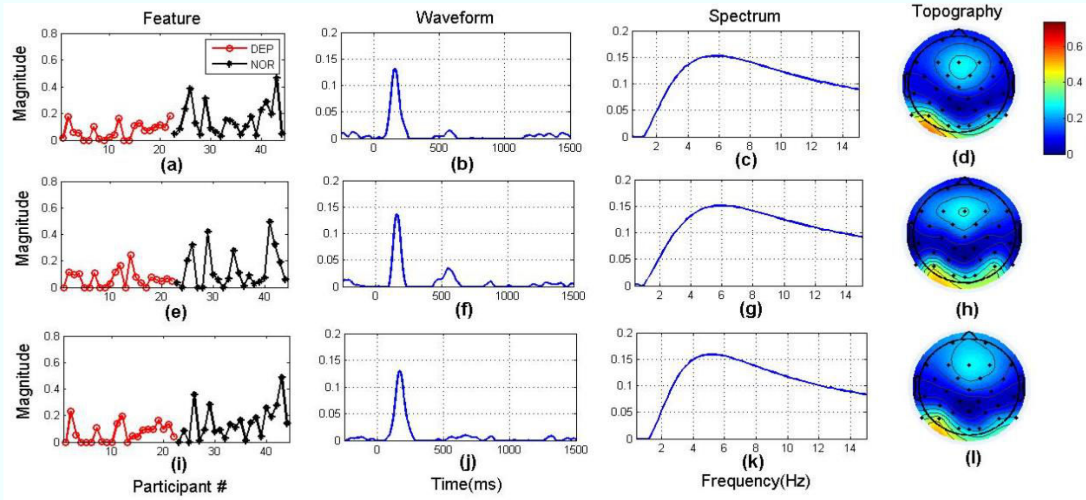

Figure 4. Feature of selected desired components in multi-domain estimated from ERPs by NTF. The first, the second, and the third rows are corresponding to the cases of happy, neutral, and sad faces respectively.

which supports the idea that the face-processing deficit in depression is not specific to emotional processing but general to faces.

V. CONCLUSION

In this study, N170 amplitude is shown to be largely specific to the faces compared with the non-face objects and unaffected by emotional facial expressions. Besides, multi-domain feature analysis of ERPs by NTF revealed that depressed patients had lower power of multi-domain feature of N170 with respect to facial stimuli, compared with normal control subjects, which may imply a perceptual early-stage dysfunction in facial information processing in depression. Moreover, depressed patients showed lower power of multi-domain feature of N170 to all types of facial stimuli, suggesting that depression do not specifically affect facial emotional perception. This study also implied that in ERP analysis multi-domain feature might be more powerful in revealing the neural dynamical difference than the conventional temporal or spectral feature.

ACKNOWLEDGMENT

The authors thank Dr. Jijun Wang for their constructive suggestions and help for the experiments, and thank Dr. Anh Huy Phan and Dr. Andrzej Cichocki in RIKEN Brain Science Institute, Japan, for providing the NTF code.

REFERENCES

[1] B. Marroquín, "Interpersonal emotion regulation as a mechanism of social support in depression,"*Clin Psychol Rev,* vol. 31, pp. 1276-1290, 2011.

[2] R. C. Gur, R. J. Erwin, R. E. Gur, A. S. Zwil, C. Heimberg, and H. C. Kraemer, "Facial emotion discrimination: II. Behavioral findings in depression," *Psychiat Res,* vol. 42, pp. 241-251, 1992.

[3] Q. Dai and Z. Feng, "More excited for negative facial expressions in depression: evidence from an event-related potential study," *Clin Neurophysiol,* vol. 123, pp. 2172-2179, 2012.

[4] S. Bentin, T. Allison, A. Puce, E. Perez, and G. McCarthy, "Electrophysiological Studies of Face Perception in Humans," *J Cognitive Neurosci,* vol. 8, pp. 551-565, 1996.

[5] M. Eimer, A. Holmes, and F. P. McGlone, "The role of spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions," *Cogn Affect Behav Neurosci,* vol. 3, pp. 97-110, 2003.

[6] M. Batty and M. J. Taylor, "Early processing of the six basic facial emotional expressions," *Cognitive Brain Res,* vol. 17, pp. 613-620, 2003.

[7] A. Cichocki, R. Zdunek, A. H. Phan, and S. Amari, *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*: Wiley, 2009.

[8] F. Cong, A. H. Phan, Q. Zhao, T. Huttunen-Scott, J. Kaartinen, T. Ristaniemi*, et al.*, "Benefits of multi-domain feature of mismatch negativity extracted by non-negative tensor factorization from EEG collected by low-density array," *Int J Neural Syst,* vol. 22, no. 1250025, 2012.

[9] F. Cong, A. Phan, P. Astikainen, Q. Zhao, J. Hietanen, T. Ristaniemi, et al., "Multi-domain feature of event-related potential extracted by nonnegative tensor factorization: 5 vs. 14 electrodes EEG data," *Lect Notes Comput Sc,* vol. 7191, pp. 502-510, 2012.

[10] N. Tottenham, J. W. Tanaka, A. C. Leon, T. McCarry, M. Nurse, T. A. Hare, et al., "The NimStim set of facial expressions: judgments from untrained research participants," *Psychiat Res,* vol. 168, pp. 242-249, 2009.

[11] S. Chen, J. Sun, and S. Tong, "Delayed attentional disengagement from sad face: A study of alpha rhythm by event-related desynchronization," *34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS),* pp. 6776-6779, 2012.

[12] M. Mørup and L. K. Hansen, "Automatic relevance determination for multi-way models," *J Chemometr,* vol. 23, pp. 352-363, 2009.

[13] R. V. Hogg and J. Ledolter, *Engineering statistics* vol. 358: MacMillan New York, 1987.

[14] B. Rossion and C. Jacques, "Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170," *Neuroimage,* vol. 39, pp. 1959-1979, 2008.

[15] J. J. Stekelenburg and B. de Gelder, "The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect," *Neuroreport,* vol. 15, pp. 777-780, 2004.

[16] A. Krombholz, F. Schaefer, and W. Boucsein, "Modification of N170 by different emotional expression of schematic faces," *Biol Psychol,* vol. 76, pp. 156-162, 2007.

[17] A. Holmes, P. Vuilleumier, and M. Eimer, "The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials," *Cognitive Brain Res,* vol. 16, pp. 174-184, 2003.

[18] S. Campanella, C. Montedoro, E. Streel, P. Verbanck, and V. Rosier, "Early visual components (P100, N170) are disrupted in chronic schizophrenic patients: an event-related potentials study," *Neurophysiol Clin,* vol. 36, pp. 71-78, 2006.