ANALYSIS OF ONGOING EEG ELICITED BY NATURAL MUSIC STIMULI USING NONNEGATIVE TENSOR FACTORIZATION

Fengyu Cong[1], Anh Huy Phan[2], Qibin Zhao[2], Asoke K. Nandi[1,3], Vinoo Alluri[4], Petri Toiviainen[4], Hanna Poikonen[5], Minna Huotilainen[5,6], Andrzej Cichocki[2], Tapani Ristaniemi[1]

1. Department of Mathematical Information Technology, University of Jyväskylä, Finland;
2. Laboratory of Advanced Brain Signal Processing, RIKEN Brain Science Institute, Japan;
3. Department of Electrical Engineering and Electronics, The University of Liverpool, UK;
4. Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyväskylä, Finland;
5. Cognitive Brain Research Unit, Institute of Behavioral Sciences, University of Helsinki, Finland;
6. Finnish Institute of Occupational Health, Helsinki, Finland

ABSTRACT

This study proposes a six-step approach to analyze ongoing EEG elicited by 512-second long natural music (modern tango). The spectrogram of the ongoing EEG was first produced, and then a fourth-order tensor including the spectrograms of multiple channels of multiple participants was decomposed via nonnegative tensor factorization (NTF) into four factors, including temporal, spectral and spatial components, and multi-domain features of all participants. We found one extracted temporal component by NTF significantly ($p < 0.01$) correlated with the temporal course of a long-term music feature, 'fluctuation centroid'; moreover, the power of posterior alpha activity was found to be associated with this temporal component. Hence, it looks promising to apply the proposed method for analyzing other ongoing EEG elicited by other natural stimuli.

Index Terms— EEG, ongoing, music, nonnegative tensor factorization, spectrogram

1. INTRODUCTION

Electroencephalography (EEG) is the recorded electrical activity along the scalp by electrodes. It was first reported by Hans Berger in 1929 [1]. At that time, the spontaneous EEG was recorded while the participant was resting, and no external stimuli were presented to the participant [1, 2]. In 1939, Davis et al. reported the event-related potentials (ERPs) which were elicited by auditory stimulations [3, 4]. ERPs are the averaged EEG activity time-locked to the presentation of repeated visual, somatosensory or auditory stimuli [4]. With development of powerful computational tools, the spontaneous EEG has been used for the clinical purposes for epilepsy, coma, tumors, stroke [2], diagnosis of brain death [5], and so on; furthermore, for the research use, the spontaneous EEG, and particularly, its derivative, ERPs, are extensively studied in the fields of neuroscience, cognitive science, psychology, physiology, and so on [4]. Moreover, the recently developed brain computer interface (BCI) indeed extends the use of spontaneous EEG and ERPs into more practical situations [6] .

Nevertheless, the spontaneous EEG is usually recorded in the rest state of a participant [2], and the ERPs are mostly collected under the specially designed presentation of stimuli [4]. Hence, these two types of data cannot straightforwardly reflect the activity in the brain in more natural and general conditions of a participant. For example, when a participant is watching a video, listening to music, moving, talking, and even playing games and sports, it will be of great interest and benefit to know the brain state of the participant through the ongoing EEG [7].

Indeed, from the view of data collection, there is no difficulty to collect the ongoing EEG data in natural conditions [7]. To the best knowledge of authors, it seems that the difficulty to study the ongoing EEG results from how to analyze the collected EEG data in such complicated, but natural, conditions. For example, how to extract components of brain activities and how to find the interesting components are still open for research. This study is devoted to this challenging topic and the ongoing EEG elicited while listening to a 512-second long piece of music will be analyzed through the proposed methodology.

Sanei and Chambers reviewed the data processing approaches for the spontaneous EEG and the ERPs [8]. For the spontaneous EEG, particularly for the data recorded when a participant is sleeping, the duration of the whole recordings can be dozens of minutes, even several hours; then, a sliding window with the length of a few seconds is often used to segment the long EEG data; next, the spectrum of the short EEG data in each segment can be analyzed; after the power spectrum at different frequency bands is calculated, the state of sleep can be concluded [2, 8]. In other words, for such long spontaneous EEG data, its spectrogram is analyzed. Regarding ERPs, the peak measurements and the event-related oscillations (EROs) [9] in the time, frequency, and time-frequency domains are often used to represent the ERP related brain activity for analysis [4, 10, 11]. Furthermore, such measurements of the spontaneous EEG and the ERPs can be used to localize the source of the brain activity when the EEG data is collected by a high-density array [2, 4, 8, 12]. In such a situation, the information of EEG data in one of time, frequency and time-frequency domains, and the spatial domain is exploited sequentially. Moreover, the multi-dimensional signal processing methods like principal/independent component analysis (PCA/ICA) and tensor factorization based on the canonical polyadic (CP) model [13] can be performed on the multi-way EEG data to exploit the information of EEG data in the time and spatial domains simultaneously [8].

From the view of data formations of spontaneous EEG, ongoing EEG and ERPs, the ongoing EEG elicited by the natural stimuli and the spontaneous EEG are very alike since the ERPs are conventionally defined as the averaged EEG over multiple single-trials, and it is very difficult to define triggers for producing single trials of EEG data of ERPs under the natural stimuli. So, referenced to the spontaneous EEG data, the spectrogram of the ongoing EEG elicited by the 512-second long music will be the object of our study, instead of the EEG waveform, hereinafter. Then, a fourth-order tensor with dimensions of time by frequency by channel by subject can be formulated via the spectrograms of ongoing EEG data of multiple channels of multiple participants. As the spectrogram is nonnegative, the nonnegative tensor factorization (NTF) based on CP model will be performed to decompose the nonnegative fourth-order tensor into the pre-defined number of components in each of four factors consisting of the temporal, spectral and spatial components, and the multi-domain features of EEG for all participants in this study [13-17]. After the decomposition through NTF, it is necessary to seek the interesting component to follow up for further analysis. This process is analogous to the application of ICA on EEG. Through ICA, a number of components can be extracted, and then, one or more interesting components are usually chosen for the further analysis based on the prior knowledge of EEG [11, 18]. In the application of NTF for the study of ERPs, the selection of the desired component is also based on the prior knowledge of the ERPs [16]. In the BCI problem regarding the spontaneous EEG, the selection of the desired components or features of EEG can be achieved through machine learning methods which are also based on prior knowledge of the spontaneous EEG [15, 19]. Hence, the prior knowledge of EEG plays an important role in selecting the interesting component for analysis.

However, regarding the ongoing EEG elicited by the natural music, we do not have enough knowledge to select the interesting component extracted by NTF from the spectrograms of EEG. Thus, the prior knowledge does not result from the EEG data, but from the music used in the experiment in this study. With the reference to the same music used in an fMRI study [20], we select the interesting component in this study through measuring the correlation coefficient between an extracted temporal component by NTF and the temporal course of a music feature. A significant correlation may indicate that the EEG data is closely associated with the music, which is of interest in the experiment. Then, spectral and spatial components parallel to the selected temporal components may reveal the spectral structure and spatial map of brain activity elicited by music.

2. METHOD

2.1. Data description

EEG data of fourteen right-handed and healthy adults aged from 20 to 46 years old were used in this study. No participants reported hearing loss or history with neurological illnesses. No participants had any musical expertise. A whole musical piece of modern tango-Astor Piazzolla was used as test stimuli, and the 8.5-minute tango of Piazzolla was recorded in a concert in Lausanne, Switzerland [20]. The EEG data was recorded according to 10-20 system with BioSemi bioactive electrode caps (64 electrodes in the cap plus 5 external electrodes at the tip of the nose, left and right mastoids and around the right eye both horizontally and vertically). The direct-current mean value between each measuring electrode and the Common Mode Sense electrode was kept under ±25 μV. EEG were collected with the sampling rate of 2048 Hz and saved for off-line processing. The external electrode of the nose was the reference and the data were preprocessed in EEGLAB [11], and then were down-sampled to 256 Hz, high-pass and low-pass filtered with 1 Hz and 30 Hz cutoff frequencies.

2.2. Nonnegative tensor factorization

In the form of tensor products, the NTF model [13] reads as

$$\underline{\mathbf{Y}} \approx \underline{\hat{\mathbf{Y}}} = \sum_{j=1}^{J} \mathbf{u}_j^{(1)} \circ \mathbf{u}_j^{(2)} \cdots \circ \mathbf{u}_j^{(N)} = \underline{\mathbf{I}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \cdots \times_N \mathbf{U}^{(N)}, \quad (1)$$

where, J is the number of extracted components, $\mathbf{u}_j^{(1)}$, $\mathbf{u}_j^{(2)} \cdots \mathbf{u}_j^{(N)}$ are associated components in N parallel factors under the CP model, $\underline{\hat{\mathbf{Y}}}$ is an approximation of the N^{th}-order tensor $\underline{\mathbf{Y}} \in \mathfrak{R}_+^{I_1 \times I_2 \times \cdots \times I_N}$, and $\underline{\mathbf{I}}$ is an identity tensor [13], $\mathbf{U}^{(n)} = \left[\mathbf{u}_1^{(n)}, \mathbf{u}_2^{(n)}, \cdots, \mathbf{u}_J^{(n)}\right] \in \mathfrak{R}_+^{I_n \times J}$ is the nonnegative matrix, $n = 1, 2, \cdots, N$, and $j = 1, 2, \cdots, J$. Most algorithms for NTF are to minimize a squared Euclidean distance via

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

In this study, we applied the hierarchical alternating least squares (HALS) algorithm [13] for NTF as

$$\mathbf{u}_j^{(n)} \leftarrow \underline{\mathbf{Y}} \,\bar{\times}_1\, \mathbf{u}_j^{(1)} \cdots \bar{\times}_{n-1}\, \mathbf{u}_j^{(n-1)} \,\bar{\times}_{n+1}\, \mathbf{u}_j^{(n+1)} \cdots \bar{\times}_N\, \mathbf{u}_j^{(N)} - \mathbf{U}_{-j}^{(n)} \left(\circledast\, {\mathbf{U}_{-j}^{(k)}}^{\mathrm{T}} \mathbf{u}_j^{k} \right), \quad (3)$$

where, '$\circledast$' denotes the Hadarmard product, $k \neq n$, and $\underline{\mathbf{Y}} \,\bar{\times}_n\, \mathbf{u}_j^{(n)}$ represents the *n*-mode product between a tensor and a vector [13]. The factors except the last one will be normalized to be unit vectors during iterations

$$\mathbf{u}_j^{(n)} \leftarrow \mathbf{u}_j^{(n)} / \left\| \mathbf{u}_j^{(n)} \right\|_2, n = 1, 2, \cdots, N-1.$$

For the decomposition of the spectrogram of ongoing EEG through NTF, we formulated a fourth-order tensor $\underline{\mathbf{Y}}$ including modes of frequency by time by channel by subject. The number of frequency bins (I_f), timestamps (I_t), channels (I_c), and subjects (I_s) compose the dimensions of the tensor $\underline{\mathbf{Y}}$. Decomposition of $\underline{\mathbf{Y}}$ results in four matrices:

$$\underline{\mathbf{Y}} \approx \underline{\mathbf{I}} \times_1 \mathbf{U}^{(f)} \times_2 \mathbf{U}^{(t)} \times_3 \mathbf{U}^{(c)} \times_4 \mathbf{F}, \quad (4)$$

where the last factor is the feature matrix ($I_s \times J$) consisting of J extracted multi-domain features of brain responses in the experiment onto the subspaces spanned by the spectral (i.e., $\mathbf{U}^{(f)}(I_f \times J)$), temporal (i.e., $\mathbf{U}^{(t)}(I_t \times J)$) and spatial (i.e., $\mathbf{U}^{(c)}(I_c \times J)$) factors. In some studies [16, 17], the multi-domain features extracted by NTF were the most interesting factor to discriminate two groups of participants. In this study, our interest lies in seeking which temporal component in the temporal factor is significantly correlated with the temporal course of a music feature, and lies in exploring the nature of those parallel spectral and spatial components associated with this temporal component. .

2.3. Number of components extracted by NTF

When the tensor factorization under the PARAFAC model is applied, a challenging problem is how to determine the number of components for extraction in each factor [13]. Three types of frequently used methods include model order selection [21], diagnosis of the consistency of core tensor named as CORCONDIA [22] in a number of tensor models, and the measurement of the change of the fit denoted as DIFFIT [23] in a number of tensor models. The former type of methods requires different assumptions for different algorithms of model order selection, and the latter two types of methods are basically data-driven. For simplicity, DIFFIT was used in this study. A fit is the explained variance of raw data by a proposed model and the fit of a NTF model is defined as

$$\mathrm{fit}(m) = 1 - \left\| \underline{\mathbf{Y}} - \underline{\widehat{\mathbf{Y}}}_m \right\|_\mathrm{F} / \left\| \underline{\mathbf{Y}} \right\|_\mathrm{F},$$

where $\underline{\widehat{\mathbf{Y}}}_m$ is the approximation for raw data tensor $\underline{\mathbf{Y}}$ in the m^{th} model, $m = 1, \cdots, M$, and $\mathrm{fit}(m)$ is monotonically rising. Then, the difference fit of the two adjacent fits is

$$\mathrm{dif}(m) = \mathrm{fit}(m) - \mathrm{fit}(m-1),$$

where $m = 2, \cdots, M$. Next, the ratio of the adjacent difference fits reads as

$$\mathrm{diffit}(m) = \mathrm{dif}(m) / \mathrm{dif}(m+1),$$

where $m = 2, \cdots, (M-1)$. The model with the largest diffit value is regarded as the appropriate NTF model for the raw data tensor $\underline{\mathbf{Y}}$ [23, 24] .

2.4. Data processing and analysis

The data processing included six steps as the following:

(**I**) Five music (tonal and rhythmic) features studied in [20] were used here. In order to extract the features, the 512-second long music was first segmented into 255 frames by a sliding rectangular window, and the duration of the frame was 3 seconds and the overlap between two adjacent frames was 1 second. The features were then extracted in each frame. Next, five temporal courses of those features were produced under the sampling frequency of 0.5 Hz. Details of the feature extraction can be found in [20, 25].

(**II**) In order to remove the artifacts caused by the eye blinks and others, a high pass digital filter was performed on the EEG with the cutoff frequency at 4 Hz. This is because the EOG activity has a wide frequency range but usually maximizes below 4 Hz [26]. In BCI studies, a high-pass filter with the cutoff frequency at 8 Hz is often used [19] for better cleaning EOG. In this study, we kept the theta oscillation of EEG for investigation, and 4 Hz was defined as the cutoff frequency. The mean value of the data at each channel was zero after the filtering.

(**III**) The short-time Fourier transform (STFT) was performed to obtain the time-frequency representation, i.e., spectrogram, of EEG. The duration of the window was four seconds, and the overlap ratio between two adjacent windows were 50%, and the number of points for Fourier transform was the same to that of the window, and the Hamming window was used. In the frequency domain, the frequency ranging from 1 to 15 Hz was used for analysis. Hence, after the STFT, a matrix was produced with the dimensions of 51 frequency bins by 255 time samples (sampling frequency was 0.5 Hz) for each channel. Here, the duration of the frame for EEG was one second longer than that of the frame for music. Such a treatment resulted from the consideration of the delay of the temporal course of the brain activity in contrast to the timing of music stimuli. When a participant listens to the auditory stimulus, the delay of the EEG can be over one hundred milliseconds [27, 28]. In this study, we did not measure the delay in the experiment; hence we defined a longer frame for EEG to reduce the delay effect to temporally match music stimuli.

(**IV**) A four-way tensor including the spectrograms of EEG was formulated and its dimensions were frequency (57 bins) by time (255 samples) by space (64 channels) by subject (14 adults), and then, NTF was performed to decompose the tensor into four factors matrices which were the spectral factor, the temporal factor, the spatial factor and the feature factor matrix for all participants, and in each factor, R components were extracted [16].

(**V**) R ranged from 6 to 65 in this study, and DIFFIT suggested when 35 components were extracted by NTF, the model of NTF for the tensor was appropriate in this study.

(**VI**) Under three significance levels with p values of 0.001 and 0.01, thresholds of correlation coefficient between the temporal course of each music feature and each of the 35 EEG temporal components in the temporal factor were investigated. Before the calculation, the mean of the temporal course was removed and normalized to its standard deviation. The method to calculate the thresholds can be found in [20]. After that, the correlation coefficient between each of the 35 EEG temporal components and each of temporal courses of five music features were calculated, the temporal components significantly correlated with the music features were then found.

3. RESULTS

Fig. 1 describes the fit of the raw tensor through the reconstructed tensor by NTF under each tested model. In terms of DIFFIT, the results of the 35-component model were selected for analysis. Fig. 2 demonstrates the correlation coefficient between each of the 35 EEG temporal components and the temporal course of the music feature, terms as fluctuation centroid (FlucCentroid), and the thresholds of the correlation coefficients for significance levels. Only the eighth component contributed a significant correlation coefficient for $p < 0.01$. Most of the extracted EEG temporal components were not significantly correlated with the temporal course of the music feature. For other music features, we did not find any EEG component significantly correlated with them under the threshold for $p < 0.01$. Fig. 3 displays the eighth EEG temporal component, and the temporal course of the music feature, 'FlucCentroid'. The correlation coefficient between them is about 0.22 which is larger than one threshold in Fig.2.

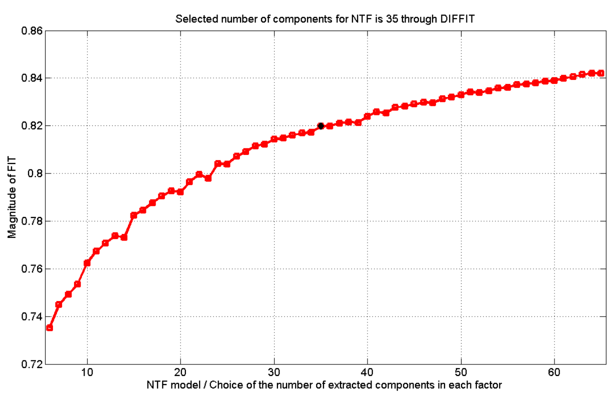

Fig. 1 Fit of every test NTF model

Fig. 4 shows the eighth spatial component, the eighth spectral component, and the eighth multi-domain feature component in the corresponding spectral factor, the spatial factor and the feature factor. The corresponding spectral component indicated that the alpha oscillation of EEG was associated with the music feature termed as 'FlucCentroid', and the parallel spatial component revealed that the posterior area was activated when participants were listening to the music regarding that feature. The 'FlucCentroid' is calculated as the geometric mean of the fluctuation spectrum representing the global repartition of rhythm periodicities within the range of 0–10 Hz.

Furthermore, the plot in the right and bottom of Fig. 4 presents the multi-domain features of EEG for all participants under the bases of the temporal, spectral and spatial EEG components shown in Fig. 3 and Fig. 4. Most participants had similar features except one participant.

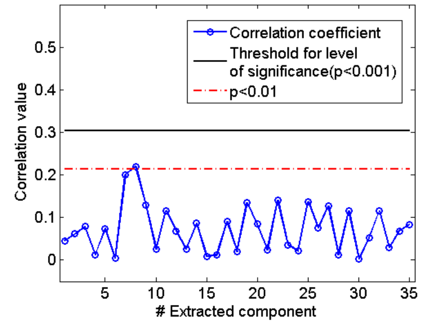

Fig. 2 Correlation coefficient between each of extracted EEG components and the temporal course of music feature-FlucCentroid

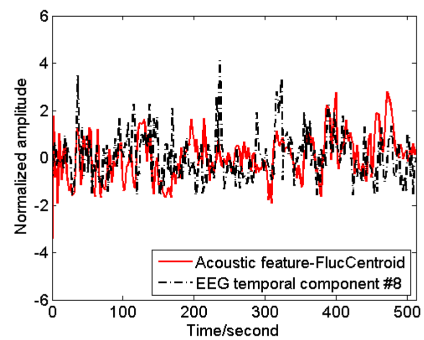

Fig. 3 Temporal courses of music feature and EEG component

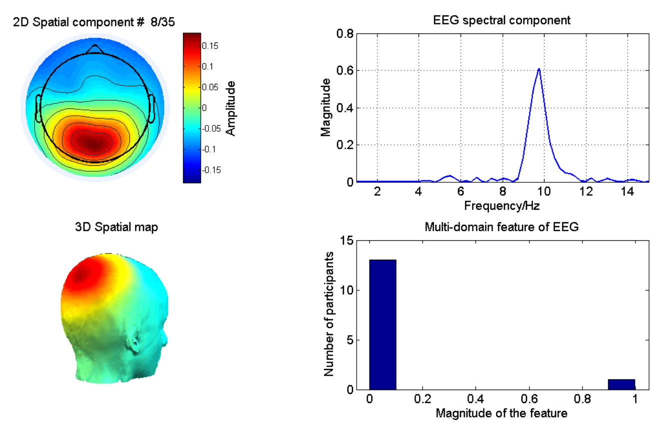

Fig. 4 Spatial component, spectral component and multi-domain feature of EEG that are parallel to the selected temporal component

4. CONCLUSION

The study addresses the decomposition of ongoing EEG elicited by natural music stimuli. The core of the proposed method is to perform NTF on the spectrograms of EEG, and to seek the interesting components with the reference of temporal courses of the music features. We found that one posterior alpha component extracted by NTF was significantly ($p < 0.01$) correlated with the temporal course of one music feature named as 'FlucCentroid' among all five long-term music features. Such result conforms to the previous discovery of music perception and imagery in EEG [29], which validates that the finding in this study does not result from the technical artifacts, but truly reveals the brain state of the participants when listening to the modern tango.

In sum, for the research of the ongoing EEG elicited by natural stimuli, it is necessary to exploit the features of the stimuli which are referenced to study the EEG. Meanwhile, such ongoing EEG can be severally contaminated by artifacts, thus, extracting meaningful EEG components and features may have to be in terms of advanced signal processing methods, namely, NTF, ICA and so on.

5. ACKNOWLEDGEMENTS

This work was financially supported by TEKES (Finland) grant 40334/10 'Machine Learning for Future Music and Learning Technologies'. AKN would like to thank TEKES for their award of the Finland Distinguished Professorship.

6. REFERENCES

[1] H. Berger, "Ueber das Elektrenkephalogramm des Menschen," *Archives Fur Psychiatrie Nervenkrankheiten,* vol. 87, pp. 527-570, 1929.

[2] E. Niedermeyer and F. Lopes da Silva, *Electroencephalography : Basic Principles, Clinical Applications, and Related Fields.* Baltimore, MD: Williams & Wilkins, 2004.

[3] H. Davis, P. A. Davis, A. L. Loomis, E. N. Harvey and G. Hobart, "Electrical reactions of the human brain to auditory stimulation during sleep," *Journal of Neurophysiology,* vol. 2, pp. 500-514, 1939.

[4] S. J. Luck, *An Introduction to the Event-Related Potential Technique.* The MIT Press, 2005.

[5] Z. Chen, J. Cao, Y. Cao, Y. Zhang, F. Gu, G. Zhu, Z. Hong, B. Wang and A. Cichocki, "An empirical EEG analysis in brain death diagnosis for adults," *Cogn. Neurodyn,* vol. 2, pp. 257-271, 2008.

[6] D. S. Tan and A. Nijholt, Eds., *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction.* London: Springer, 2010.

[7] S. Makeig, K. Gramann, T. P. Jung, T. J. Sejnowski and H. Poizner, "Linking brain, mind and behavior," *Int. J. Psychophysiol.,* vol. 73, pp. 95-100, Aug, 2009.

[8] S. Sanei and J. A. Chambers, *EEG Signal Processing.* Wiley, 2007.

[9] G. Y. Görsev and E. Basar, "Sensory evoked and event related oscillations in Alzheimer's disease: a short review," *Cogn Neurodyn,* vol. 4, pp. 263-274, 2010.

[10] E. Basar, M. Schurmann, T. Demiralp, C. Basar-Eroglu and A. Ademoglu, "Event-related oscillations are 'real brain responses'--wavelet analysis and new strategies," *Int. J. Psychophysiol.,* vol. 39, pp. 91-127, Jan, 2001.

[11] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods,* vol. 134, pp. 9-21, Mar 15, 2004.

[12] L. Fuentemilla, J. Marco-Pallares, T. F. Munte and C. Grau, "Theta EEG oscillatory activity and auditory change detection," *Brain Res.,* vol. 1220, pp. 93-101, Jul 18, 2008.

[13] A. Cichocki, R. Zdunek, A. H. Phan and S. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis.* John Wiley, 2009.

[14] A. Cichocki, Y. Washizawa, T. M. Rutkowski, H. Bakardjian, A. H. Phan, S. Choi and Q. Zhao, "Noninvasive BCIs: Multiway Signal-Processing Array Decompositions," *Computer,* vol. 41, pp. 34-42, 2008.

[15] H. Lee, Y. D. Kim, A. Cichocki and S. Choi, "Nonnegative tensor factorization for continuous EEG classification," *Int. J. Neural Syst.,* vol. 17, pp. 305-317, Aug, 2007.

[16] F. Cong, A. H. Phan, P. Astikainen, Q. Zhao, J. K. Hietanen, T. Ristaniemi and A. Cichocki, "Multi-domain Feature of Event-related Potential Extracted by Nonnegative Tensor Factorization: 5 vs.14 Electrodes EEG Data," *Lecture Notes in Computer Science (F. Theis et al. (eds.): LVA/ICA 2012),* vol. 7191, pp. 502-510, 2012.

[17] F. Cong, A. H. Phan, H. Lyytinen, T. Ristaniemi and A. Cichocki, "Classifying Healthy Children and Children with Attention Deficit through Features Derived from Sparse and Nonnegative Tensor Factorization Using Event-related Potential," *Lecture Notes in Computer Science (in V. Vigneron et al. (eds.): LVA/ICA 2010),* vol. 6365, pp. 620-628, 2010.

[18] F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X. Huang, H. Lyytinen and T. Ristaniemi, "Answering Six Questions in Extracting Children's Mismatch Negativity through Combining Wavelet Decomposition and Independent Component Analysis," *Cogn. Neurodynamics,* vol. 5, pp. 343-359, 2011.

[19] Q. Zhao, T. M. Rutkowski, L. Zhang and A. Cichocki, "Generalized optimal spatial filtering using a kernel approach with application to EEG classification," *Cogn. Neurodyn,* vol. 4, pp. 355-358, Dec, 2010.

[20] V. Alluri, P. Toiviainen, I. P. Jaaskelainen, E. Glerean, M. Sams and E. Brattico, "Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm," *Neuroimage,* vol. 59, pp. 3677-3689, 2012.

[21] Z. He, A. Cichocki, S. Xie and K. Choi, "Detecting the number of clusters in n-way probabilistic clustering," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 32, pp. 2006-2021, Nov, 2010.

[22] R. Bro and H. A. L. Kiers, "A new efficient method for determining the number of components in PARAFAC models," *Journal of Chemometrics,* vol. 17, pp. 274-286, 2003.

[23] M. E. Timmerman and H. A. Kiers, "Three-mode principal components analysis: choosing the numbers of components and sensitivity to local optima," *Br. J. Math. Stat. Psychol.,* vol. 53 (Pt 1), pp. 1-16, May, 2000.

[24] M. Mørup and L. K. Hansen, "Automatic Relevance Determination for Multiway Models," *Journal of Chemometrics,* vol. 23, pp. 352-363, 2009.

[25] O. Lartillot and P. Toiviainen , "MIR in matlab (II): A toolbox for musical feature extraction from audio," in 2007, pp. 237-244.

[26] M. Fatourechi, A. Bashashati, R. K. Ward and G. E. Birch, "EMG and EOG artifacts in brain computer interface systems: A survey," *Clin. Neurophysiol.,* vol. 118, pp. 480-494, Mar, 2007.

[27] D. A. Abrams, T. Nicol, S. Zecker and N. Kraus, "Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech," *J. Neurosci.,* vol. 28, pp. 3958-3965, 2008.

[28] S. J. Aiken and T. W. Picton, "Envelope and spectral frequency-following responses to vowel sounds," *Hear. Res.,* vol. 245, pp. 35-47, Nov, 2008.

[29] R. S. Schaefer, R. J. Vlek and P. Desain, "Music perception and imagery in EEG: alpha band effects of task and stimulus," *Int. J. Psychophysiol.,* vol. 82, pp. 254-259, Dec, 2011.