Linking Brain Responses to Naturalistic Music through Analysis of Ongoing EEG and Stimulus Features

Fengyu Cong, *Member, IEEE*, Vinoo Alluri, Asoke K. Nandi, *Fellow, IEEE*, Petri Toiviainen, Rui Fa, *Member, IEEE*, Basel Abu-Jamous, *Member, IEEE*, Li-yun Gong, Bart G. W. Craenen, Hanna Poikonen, Minna Huotilainen, Tapani Ristaniemi, *Senior Member, IEEE*

***Abstract*—This study proposes a novel approach for the analysis of brain responses in the modality of ongoing EEG elicited by the naturalistic and continuous music stimulus. The 512-second long EEG data (recorded with 64 electrodes) are first decomposed into 64 components by independent component analysis (ICA) for each participant. Then, the spatial maps showing dipolar brain activity are selected in terms of the residual dipole variance through a single dipole model in brain imaging, and clustered into a pre-defined number (estimated by the minimum description length) of clusters. Subsequently, the temporal courses of the EEG theta and alpha oscillations of each component for each cluster are produced and correlated with the temporal courses of tonal and rhythmic features of the music. Using this approach, we found that the extracted temporal courses of the theta and alpha oscillations along central and occipital area of scalp in two of the selected clusters significantly correlated with the musical features representing progressions in the rhythmic content of the stimulus. We argue that this demonstrates that with the proposed approach we have managed to discover what kinds of brain responses were elicited when a participant was listening continuously to the long piece of naturalistic music.**

***Index Terms*—Acoustical features, Clustering, EEG, Independent component analysis, Natural continuous music, Ongoing, Oscillation**

Manuscript received April 16, 2012.

Fengyu Cong, Vinoo Alluri, Asoke K. Nandi and Tapani Ristaniemi are with Department of Mathematical Information Technology, University of Jyväskylä, Finland, E-mail: FengyuCong@gmail.com, vinoo.alluri@jyu.fi, asoke.nandi@brunel.ac.uk, tapani.ristaniemi@jyu.fi

Vinoo Alluri and Petri Toiviainen are with Finnish Centre of Excellence in Interdisciplinary Music Research, Department of Music, University of Jyväskylä, E-mail: vinoo.alluri@jyu.fi, petri.toiviainen@jyu.fi

Asoke K. Nandi, Fa Rui, Basel Abu-Jamous, and Bart G. W. Craenen are with Department of Electronic and Computer Engineering, Brunel University, Uxbridge, Middlesex, UB8 3PH, UK, Email: asoke.nandi@brunel.ac.uk, Rui.Fa@brunel.ac.uk, baselabujamous@ieee.org, bcraenen@xs4all.nl

Li-yun Gong is with Department of Electrical Engineering and Electronics, The University of Liverpool, UK, Email: l.gong@liverpool.ac.uk,

Hanna Poikonen is with Cognitive Brain Research Unit, Institute of Behavioral Sciences, University of Helsinki, Finland, E-mail: hapoikon@mappi.helsinki.fi

Minna Huotilainen is with Finnish Institute of Occupational Health, Helsinki, Finland and Cognitive Brain Research Unit, Institute of Behavioral Sciences, University of Helsinki, E-mail: minna.huotilainen@helsinki.fi

I. Introduction

REVEALING brain states during real-word experiences has been an attractive problem in the past few decades. However, due to the complexity of the human brain, the state under the naturalistic stimuli including music and video has only recently been decoded through functional magnetic resonance imaging (fMRI) [1-6] and magnetoencephalography (MEG) [7, 8]. Brain states during real-world experiences, resulting in relatively low signal-to-noise ratio (SNR) in collected data, are in general more complicated to analyze, than those recorded during the resting state or under the controlled and rapidly repeated stimuli. FMRI and MEG possess better properties for source localization than another well-known method for brain imaging: electroencephalography (EEG) [9]. Furthermore, in the order of fMRI, MEG and EEG, the SNR in the corresponding datasets tends to decrease. As a result, the brain imaging during the real-world experiences has been mostly studied through fMRI. However, due to the surprisingly high cost in establishing and maintaining relevant laboratories, fMRI and MEG are not as extensively used as EEG, which does limit the extensive study of brain imaging under real-word experiences. Previous studies have shown that the listener's fMRI data can be significantly correlated with the music stimulus [1] and that from the listener's MEG signals fragments of naturalistic speech can be identified [8]. Hence, it would be natural to infer that the listener's ongoing EEG data can be closely associated with an auditory stimulus including speech and music as well. As such, this study is targeted to formulate an approach for linking brain responses to naturalistic and continuous music by analyzing the elicited ongoing EEG and correlating it to the music and its features. Using this approach, the authors hope that brain states during real-word experiences may be investigated more extensively in more brain imaging laboratories.

EEG is the recorded electrical activity along the scalp by electrodes. It was first reported by Hans Berger in 1929 [10]. At that time, spontaneous EEG was recorded while a participant was resting, and no external stimuli were presented to the participant [10, 11]. In 1939, Davis et al. reported event-related potentials (ERPs) which were elicited by auditory stimuli [12, 13]. ERPs are the averaged EEG activity time-locked to the presentation of repeated visual, somatosensory or auditory stimuli [13]. With the development of powerful computational tools, spontaneous EEG have been used for the clinical purposes such as epilepsy, coma, tumors, stroke [11], diagnosis of brain death [14], and so on. Furthermore, spontaneous EEG, and particularly, its derivative, ERPs, are extensively studied in the fields of neuroscience, psychology, physiology, and cognition [13]. Moreover, the recently developed brain computer interface (BCI) extends the use of spontaneous EEG and ERPs in more practical situations [15]. Nevertheless, spontaneous EEG is still usually recorded with the participant in a resting state [11], and the ERPs are mostly collected under the specially designed presentation of controlled stimuli [13]. Hence, these two types of EEG data cannot straightforwardly reflect the activity in the brain in more naturalistic and general conditions of a participant, for example, when a participant is watching a video, listening to music, even walking [16]. It is of great interest and benefit to know the brain state of a participant through ongoing EEG [17].

One of the difficulties of studying ongoing EEG lies in analyzing such data in the complicated and more naturalistic and continuous conditions. The question of how to decompose the data and how to find components of interest still remains open for scientific research. This study will focus on the challenging topic of examining elicited ongoing EEG collected while the participant was listening to a 512-second long piece of continuous music. A novel method for analyzing ongoing EEG data is proposed.

II. Overview

Sanei and Chambers reviewed data processing approaches for spontaneous EEG and ERPs [18]. For spontaneous EEG, particularly for data recorded when a participant is sleeping, the duration of the whole recording 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 power spectrum of the short EEG data in each segment can be analyzed based on EEG oscillations with different frequency bands. Subsequently, the state of sleep can be concluded [11, 18]. In other words, for such long spontaneous EEG data, EEG oscillations based on spectrogram of the data are analyzed.

For ERPs, the peak measurements based on the stimulus onset and the event-related oscillations (EROs) [19] in the time, frequency, and time-frequency domains are often used to represent the ERP related brain activity for analysis [13, 20, 21]. 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, in order to simultaneously exploit the information of EEG data in the time and spatial domains, multi-dimensional signal processing methods consisting of principal or/and independent component analysis (PCA/ICA) and tensor factorization can be performed on the multi-way EEG data [18].

From the view of data formations of spontaneous EEG, ongoing EEG, and ERPs, ongoing EEG elicited by the naturalistic and continuous stimulus actually possess the continuity property of spontaneous EEG and the event-related characteristic of ERPs. Hence, the analysis of ongoing EEG can be based on the EEG oscillations elicited by the naturalistic and continuous stimulus. Indeed, EEG oscillations represent cognitive functions [22]. Particularly, theta and alpha have been elicited by the controlled and repeated music pieces which were short [23-25]. When the music stimulus is naturalistic, continuous and long, the collected ongoing EEG can be segmented first, and then, the power of theta and alpha can be calculated to formulate the temporal courses of EEG oscillations for further investigation. We follow this idea here.

Collected EEG data are generally assumed to be linear mixtures of a number of unobserved and underlying electrical brain activities [26, 27]. It is necessary to separate the mixtures, i.e., the recorded EEG data, to obtain the desired underlying brain activities (EEG oscillations elicited by the naturalistic and continuous music stimulus in this study). In order to achieve this goal, filters, such as, a digital filter and wavelet filter [28, 29] with optimal passbands, and spatial filters like ICA [21, 30], are often used to filter the collected EEG, as well as a combination of them [31, 32]. In this study, we used ICA to decompose 64-channel ongoing EEG into 64 independent components for each participant.

However, with ongoing EEG elicited by the naturalistic and continuous music, we do not have enough knowledge from such ongoing EEG to select the component of interest extracted by ICA. Thus, in the experiment described here, prior knowledge does not originate from the EEG data, but from the music used as the stimulus. We select relevant components by measuring the correlation coefficients between temporal courses of theta and alpha oscillation of a temporal component extracted by ICA and the temporal course of each musical feature. A significant correlation indicates that the EEG data is closely associated with the music, which is of interest in the experiment. Then, the spatial component parallel to a selected temporal component reveals the spatial map of the brain activity elicited by the music.

The remainder of the study is then structured as follows: Methods for ongoing EEG data processing and analysis are introduced in Section III, with the results presented in Section IV. Finally, the discussion and conclusion on the results and the methods are presented in Section V.

III. Method

A. Data description

This study uses the EEG data of fourteen right-handed and healthy adults aged 20 to 46 years old. No participants reported hearing loss or had history of neurological illnesses. None of participants had musical expertise. During the experiment, participants were told to listen to music and sit as still as possible with eyes open. An 8.5-minute long musical piece of modern tango by Astor Piazzolla was used as the stimulus [1]. The EEG data were recorded with 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 used as the reference and the data were preprocessed in EEGLAB [21], and then were down-sampled to 256 Hz, and high-pass and low-pass filtered with 1 Hz and 30 Hz cutoff frequencies.

After the experiment, the EEG data were visually checked and no obvious artifacts from head movement were found. The data were used for further analysis.

B. Musical features

Five musical (tonal and rhythmic) features studied in [1] are examined here. The features were extracted using a frame-by-frame analysis approach commonly used in the field of Music Information Retrieval (MIR). The duration of each frame was 3 seconds and the overlap between two adjacent frames was 1 second, thereby resulting in five musical feature temporal courses at a sampling frequency of 0.5 Hz. All features were extracted using the MIR toolbox [1, 33].

For completeness, the five features are briefly introduced below. Two tonal and three rhythmic features were extracted. For the former, Mode, i.e., strength of major of minor mode, and, Key Clarity, i.e., the measure of the tonal clarity, were produced. The rhythmic features include Fluctuation Centroid, Fluctuation Entropy, and Pulse Clarity [1]. Fluctuation centroid is the geometric mean of the fluctuation spectrum representing the global repartition of rhythm periodicities within the range of 0–10 Hz [1]. This feature indicates the average frequency of these periodicities. Fluctuation entropy is the Shannon entropy of the fluctuation spectrum representing the global repartition of rhythm periodicities. It is a measure of the noisiness of the fluctuation spectrum. For example, a noisy fluctuation spectrum can be indicative of several co-existing rhythms of different periodicities, thereby indicating a high level of rhythmic complexity [1]. Pulse Clarity, naturally, is an estimate of clarity of the pulse [1].

C. Conventional data analysis based on spectrogram in the electrode field

The short-time Fourier transform (STFT) was applied to the filtered EEG in order to obtain the corresponding spectrogram, i.e., time-frequency representation. The duration of the window was three seconds, the overlap ratio between two adjacent windows was 33.3%, and the number of points for Fourier transform was 1024, and in addition a Hamming window was used. After the spectrogram of the EEG data was obtained, the temporal course of an EEG oscillation can be produced by integrating the power of the spectrogram over the frequency range of the oscillation at each timestamp of the spectrogram. For example, the theta oscillation ranges from 4 to 8 Hz and the alpha is from 8 to 13 Hz [11]. Subsequently, the temporal courses of theta and alpha oscillations were produced at each channel. Then, each temporal course of an EEG oscillation was correlated with each temporal course of each musical feature. Hence, for a pair of one EEG oscillation and one musical feature, there was one correlation coefficient at each electrode.

D. Independent component analysis

ICA has been extensively used to study brain signals [30], and it is based on the linear transformation model associating the EEG recordings ($\mathbf{x}$) along the scalp and the electrical sources ($\mathbf{s}$) in the brain. The model without sensor noise can be expressed as

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{1}$$

where $\mathbf{x} = [x_1, x_2, \cdots, x_I]^T$, $\mathbf{s} = \left[s_1, s_2, \cdots, s_J\right]^T$, and $\mathbf{A}$ with the full column rank is usually called as the mixing matrix regarding ICA, and in this study we designate it as the mapping matrix containing coefficients to map sources in the brain to points along the scalp. For any source, its mapping can be illustrated as

$$\mathbf{x}_r = \mathbf{a}_r \cdot s_r, \tag{2}$$

where, $\mathbf{x}_r = \left[x_{1,r}, x_{2,r}, \cdots, x_{I,r}\right]^T$, $\mathbf{a}_r$ is one column of $\mathbf{A}$ with $r \in [1, J]$. In this case, $\mathbf{x}_r$ is not the mixture like $\mathbf{x}$ in (1) any more, but is the sole information of one brain source. Hence, one goal to apply ICA is to extract the mapping of one source in (2) from the mixture in (1) [34, 35]. For simplicity without losing generality, we assume $I = J$ here. In order to obtain (2), an unmixing matrix is first learned by ICA [36], and then it transforms the mixture in (1) into independent components as

$$\mathbf{y} = \mathbf{W}\mathbf{x}. \tag{3}$$

Usually, any component of interest is selected according to prior knowledge from EEG and/or from the stimulus and is then projected back to the electrode field to correct the variance indeterminacy of the extracted component by ICA [26, 27, 37-39] through

$$\mathbf{e}_k = \mathbf{b}_k \cdot y_k, \tag{4}$$

$$\mathbf{B} = \mathbf{W}^{-1}, \tag{5}$$

where $\mathbf{e}_k = \left[e_{1,k}, e_{2,k}, \cdots, e_{I,k}\right]^T$, $\mathbf{b}_k$ is one column of $\mathbf{B}$, y_k is one element of $\mathbf{y}$ and $k \in [1, I]$. In this way, we obtain the desired electrical brain activity's magnitude with the unit of the microvolt in the context of EEG recordings [26, 27]. Furthermore, the global matrix of ICA can be defined as [40],

$$\mathbf{C} = \mathbf{W}\mathbf{A}, \tag{6}$$

where c_{ki} is the (k, i) element of the global matrix $\mathbf{C}$. Only under the perfect ICA decomposition (refereed as global optimization in [34, 35]), there is only one nonzero element in each row and column of $\mathbf{C}$ [40], and then, Eq. (4) turns [34, 35]

$$\mathbf{e}_k = \mathbf{b}_k \cdot (c_{kr} s_r) \overset{\mathbf{BC=A}}{=} \mathbf{a}_r \cdot s_r = \mathbf{x}_r, \tag{7}$$

where c_{kr} is the nonzero element.

Under the perfect ICA decomposition, we obtain the following points: 1) the estimated $\mathbf{b}_k$ is the scaled version of the mapping coefficients $\mathbf{a}_r$, i.e., the spatial map of the relevant source, 2) y_k is the scaled version of s_r, and 3) $\mathbf{e}_k$, i.e., projection of y_k, is equal to $\mathbf{x}_r$, i.e., the mapping of s_r along the scalp [34, 35]. So, the projection of an ICA component in the electrode field does not have variance or polarity indeterminacy under the perfect ICA decomposition in theory [34, 35]. For example, ICA is run twice on the same dataset and the perfect ICA decomposition is obtained in each round. If y_{k_1} and y_{k_2} are associated with the source s_r in the respective two rounds, the absolute value of correlation coefficient between y_{k_1} and s_r, or between y_{k_2} and s_r, is '1', and y_{k_1} and y_{k_2} tend to be different scaled versions of s_r, and their projections in the electrode field are completely identical and the same as $\mathbf{x}_r$.

However, practically, the perfect ICA decomposition is hard to obtain, which means that there are more than one nonzero element in some rows and column of $\mathbf{C}$ [34, 35]. In this case, $\mathbf{b}_k$, y_k and $\mathbf{e}_k$ just approximate the scaled versions of $\mathbf{a}_r$, s_r and $\mathbf{x}_r$, respectively, and the errors are associated with the performance of ICA decomposition, i.e., how well the mixtures are separated by ICA [34, 35]. For example, ICA is run twice on the same dataset and the imperfect ICA decomposition is obtained in each round. If y_{k_1} and y_{k_2} in the two rounds are regarded as the components closely associated with the source s_r, the absolute value of correlation coefficient between y_{k_1} and s_r, or between y_{k_2} and s_r, or between y_{k_1} and y_{k_2}, is smaller than '1', and projections of y_{k_1} and y_{k_2} in the electrode field are probably not completely identical. Therefore, the stability of ICA decomposition is one of the critical issues when ICA is used to study brain signals [30].

An evaluation of the performance of ICA decomposition for our study will be given in Results section. In this study, unlike the analysis of $\mathbf{e}_k$, i.e., the multiplication between $\mathbf{b}_k$ and y_k, in the research of ERPs [21, 26, 27, 31, 41], we will exploit $\mathbf{b}_k$ and y_k in the component space hereinafter. The stability and reliability of the ICA components will be analyzed next to confirm that the ICA decomposition is acceptable for further analysis to draw any conclusions.

E. ICA decomposition-stability and reliability

ICA was applied to the EEG data to estimate the unmixing matrix and extract 64 independent components for every participant using ICASSO software [42]. The reason to extract 64 components resulted from 64 electrodes used to collect the EEG data [21]. For ICASSO, InfomaxICA [43] was applied 30 times and 64 components were extracted at each run. Each time the default set for other parameters and algorithms was used.

The advantage of this software is that the stability of ICA decomposition can be analyzed. It may run one ICA algorithm many times respectively with individually and randomly initialized unmixing matrices; then, all the extracted components are clustered into the predefined number of clusters; finally, each common component in each cluster represents one component extracted by ICASSO and the stability index denoted by IQ is calculated for such a component [42]. IQ is the cluster quality index to reflect the compactness and isolation of a cluster as

$$\mathrm{IQ} = \bar{S}(i)_{int} - \bar{S}(i)_{ext}$$

where $\bar{S}(i)_{int}$ and $\bar{S}(i)_{ext}$ are averages of intra- and extra-cluster similarities [42], respectively, and $i = 1, \cdots, J$, where J is the number of clusters. The IQ ranges from '0' to '1'. When IQ approaches '1', it means that the corresponding component is extracted in almost every ICA decomposition application. This indicates a high stability of the ICA decomposition for that component. Otherwise, it means the ICA decomposition is not stable. Correspondingly, if all the clusters are isolated with each other, ICA decomposition should be stable and satisfactory. Otherwise, the extracted components are unacceptable for further analysis. We use magnitude of IQ as the criterion to evaluate stability of the ICA decomposition in this study.

ICASSO just reports the stability of ICA decomposition. In

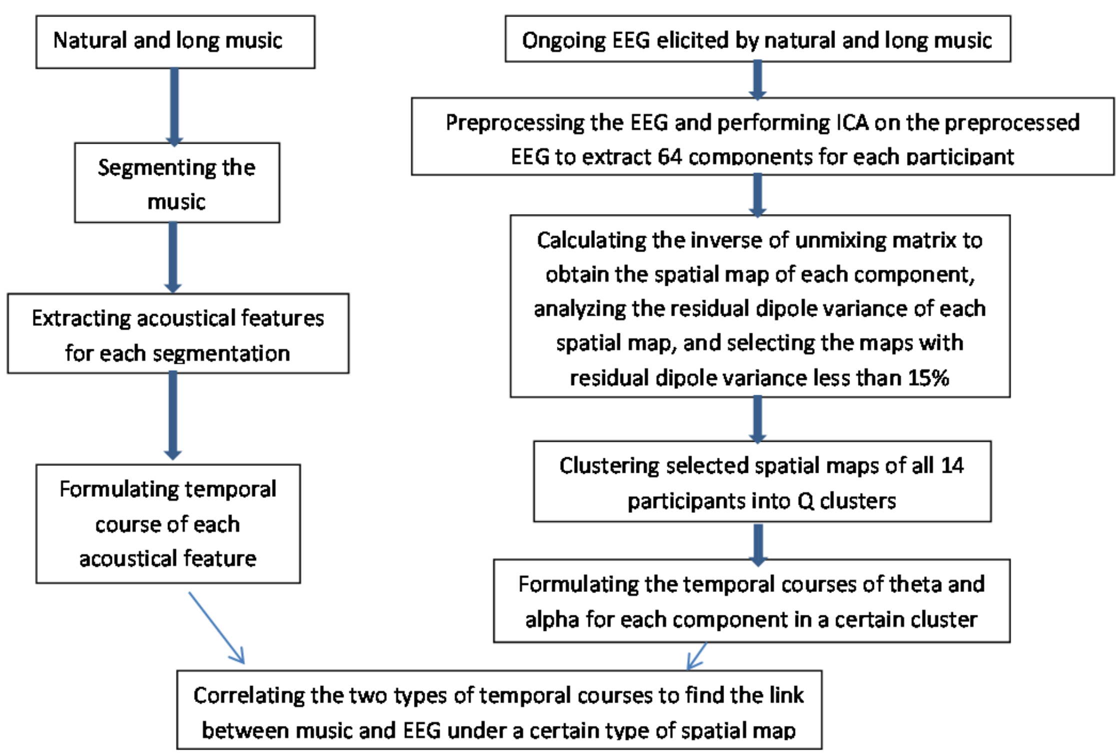

Fig. 1. Diagram of advanced data processing and analysis for ongoing EEG

order to validate the reliability of an ICA component, we follow the approach analyzing the residual dipole variance of each component through a single equivalent dipole model in brain imaging [44]. It has been recently found that independent EEG sources are dipolar [44]. As a result, if most of the ICA components are not dipolar, ICA decomposition is not acceptable even if its stability is high.

After the analysis of residual dipole variance of each component, the components with the residual dipole variance less than 15% were chosen for the further analysis.

F. Clustering selected spatial maps showing dipolar activity

In this study, each component has one spatial map with 64 parameters due to 64 electrodes. Since ICA was performed on individual EEG datasets, we can find the common information among most of participants by clustering spatial maps in order to draw a reliable conclusion. For reliable clustering, the mean of each spatial map was equalized and each map was normalized towards its standard deviation. All the selected spatial maps showing the dipolar brain activity were clustered into Q clusters to find the common spatial maps across most of participants. For clustering, we used '*k*-means' clustering algorithm [45] with the Kaufman Approach (KA) [46] for initializing the algorithm. The number of clusters, Q, was determined by a model order selection method called the minimum description length (MDL) [47].

During real-world experiences, finding relevant brain activities elicited by naturalistic and continuous auditory stimulus can be realized by correlating the temporal courses of brain activities and the temporal courses of features of the stimulus [1, 8]. If the correlation is significant, the corresponding brain activity is regarded to be associated with the stimulus [1, 8]. Fig.1 summarizes the advanced data processing and analysis for ongoing EEG in this study.

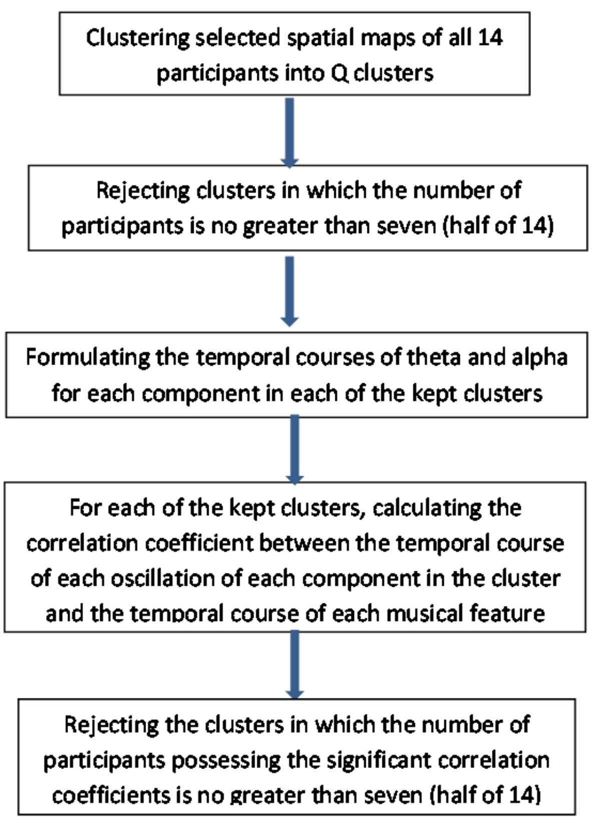

Fig. 2. Diagram to find interesting clusters

G. Finding contributors to interesting clusters

In order to find information in the ongoing EEG shared across the majority of participants, after clustering, we checked, for each cluster, the number of participants whose theta or alpha oscillation was correlated with at least one of the five musical features. If the number of participants in one cluster was less than half of all participants, that cluster would be rejected for further analysis. Mostly, this is because we consider that such a cluster does not reveal information shared among enough participants. Subsequently, the temporal courses of theta and alpha oscillations were calculated based on the spectrogram of the ICA component corresponding to each spatial map in each kept cluster. The parameters to calculate the spectrogram of an ICA component are the same as those for conventional analysis.

For each kept cluster, correlation analysis was performed between the temporal course of each EEG oscillation and that of each musical feature. As long as a participant in a kept cluster possesses one significant correlation coefficient between any oscillation and any musical feature, the brain responses of this participant are considered to be associated with the music stimulus. Then, this participant is regarded as one effective contributor to the cluster. If there are more than half of participants who are effective contributors in one cluster, this cluster is concluded to represent the common spatial maps associated with the music stimulus across most of participants. Also, such a cluster is the cluster of interest in this study for conclusions. Fig.2 summarizes all the steps.

After one cluster was chosen, the included spatial maps, possessing at least a significant correlation coefficient between theta or alpha oscillation and one of the five musical features, were paralleled and averaged to produce the final spatial map of the selected cluster.

H. Determining level of significance of correlation

Statistically, it is necessary to investigate the significance of the correlation coefficient between two temporal courses [1]. Only those components whose temporal courses of brain responses significantly correlate with the temporal courses of musical features would be considered to be relevant to this study. Then, the threshold to determine the significant correlation coefficient should be given. For the conventional analysis of ongoing EEG in this study, one threshold was based on one musical feature and 896 temporal courses (64 electrodes by 14 participants) of an EEG oscillation. For the advanced data analysis, one threshold was derived from one musical feature and K (K is the number of spatial maps in one cluster) temporal courses of an oscillation. Since the temporal courses of EEG oscillations and musical features are inevitably correlated many times, correction for multiple comparisons should be applied to counter the reduction in statistical power [48]. For this purpose, the Monte-Carlo method presented in [1] and the permutation test procedure [48] were employed to calculate the significance of correlation coefficient and to correct for multiple comparisons. Given a pair of one musical feature and one oscillation, one threshold to determine the significant correlation coefficient was calculated. Pearson correlation analysis was applied. The components of interest were then determined as the ones displaying significant correlations ($p < 0.05$) with the musical features.

By such correlation analysis, for the conventional data analysis approach, the temporal course of the EEG oscillation activated by the music stimulus could be found in the electrode field along the scalp. For the advanced data analysis, the interesting spatial map extracted by ICA was determined based on the music-associated temporal courses of the EEG oscillations of the ICA components in a cluster. In this way, the brain responses in the time, frequency and spatial domains can be linked to the music stimulus [1].

IV. Results

A. Musical features

Fig. 3 shows the temporal courses of five musical features. The coherences between the Pulse Clarity and Fluctuation Entropy, between Fluctuation Entropy and Fluctuation Centroid, and between Mode and Key are significant ($p < 0.05$). For other pairs of musical features, the correlation is not significant.

B.

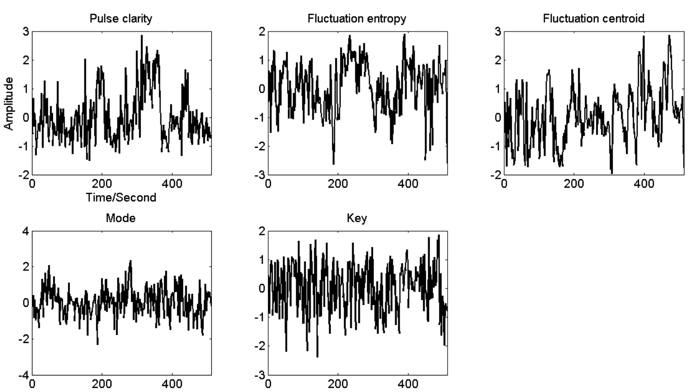

Fig. 3. Temporal courses of five musical features.

C. Conventional ongoing EEG data analysis

The ongoing EEG data were 512 seconds long which makes it difficult to present the waveform of the ongoing EEG as shown in most of EEG related studies [11, 13]. In order to describe the recorded data, the spectrograms of the grand averaged ongoing EEG over all the 14 participants at four channels are shown in Fig. 4. It can be seen that the alpha oscillation was relatively evident at Pz and Oz, and that the alpha and the theta appeared at Cz. Indeed, this is based on the grand averaged EEG and no information of the individual participants can be exploited.

Fig. 5 shows the absolute value of the correlation coefficients between the temporal courses of two oscillations and the temporal courses of five musical features at each electrode. Different curves in the figure represent different participants, and the dashed straight line denotes the threshold relating to the significance level with $p < 0.05$. For majority of participants, the correlation is not significant. Hence, it is difficult to identify the brain activities associated with the music stimulus by the conventional analysis. Therefore, this motivates us to analyze the ongoing EEG with advanced signal processing techniques.

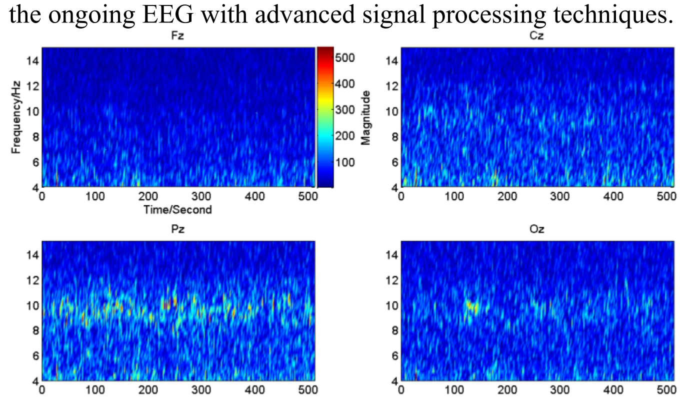

Fig. 4. Spectrograms (4-15 Hz) of the grand average of raw ongoing EEG data. The color from 'blue' to 'red' represents rising magnitude.

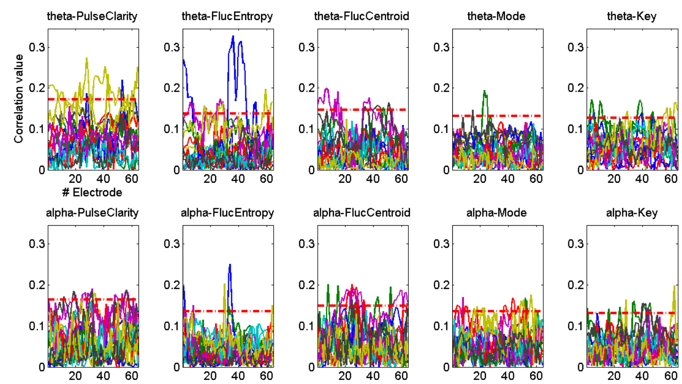

Fig. 5. Absolute values of correlation coefficients between the temporal courses of five musical features and the temporal courses of alpha oscillations of the preprocessed ongoing EEG (bottom) at each electrode as well as absolute values of correlation coefficients between the temporal courses of five musical features at each electrode and the temporal courses of theta oscillations (top). Different curves in the figure represent different participants, and the dashed straight line denotes the threshold at the significance level with $p < 0.05$.

D. ICA decomposition-Stability and Reliability

In terms of IQ given by ICASSO in Fig. 6, ICA decomposition in this study was stable and satisfactory for all the participants. For most components, their IQs were greater than 0.75. From the view of clustering in ICASSO, the 64 clusters were isolated with each other for each participant, indicating ICA decomposition was stable.

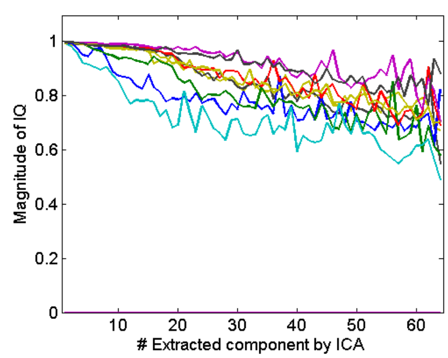

Fig. 6. IQ of each component extracted by ICA through ICASSO. Different curves represent different participants.

Fig. 7 describes the histogram of the number of ICA components among 896 ICA components under each possible residual dipole variance. The number of components with the residual dipole variance no greater than 15% is 304 among all 896 components (about 34.0%). Such parameters indicate an acceptable ICA decomposition in this study since quite many components are dipolar.

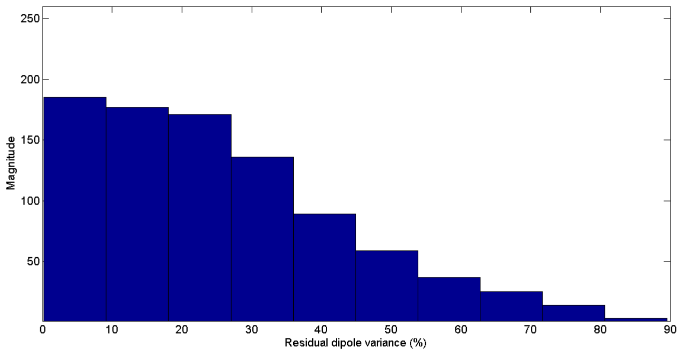

Fig. 7. The number of ICA components among 896 ICA components under one possible residual dipole variance.

E. Interesting clusters

After 304 spatial maps with residual dipole variance less than 15% were chosen, MDL was performed to estimate the number of clusters as nine. The spatial maps were then clustered into nine clusters. After the contributors were found in each cluster, clusters #1 and #5 were chosen as only they satisfied the criteria for a cluster of interest. Fig. 8 shows the spatial maps of the clusters #1 and #5, revealing that the ongoing EEG along the central and occipital area in this study could be relevant to the music stimulus. Moreover, regarding cluster #1, the mean correlation coefficient between individual spatial maps in this cluster and the spatial map of this cluster (i.e., the average of the individual spatial maps) is 0.81, and the corresponding standard deviation (SD) is 0.11. For cluster# 5, the mean is 0.85, and SD is 0.11. This indicates the memberships in cluster #1 or #5 are highly correlated with each other.

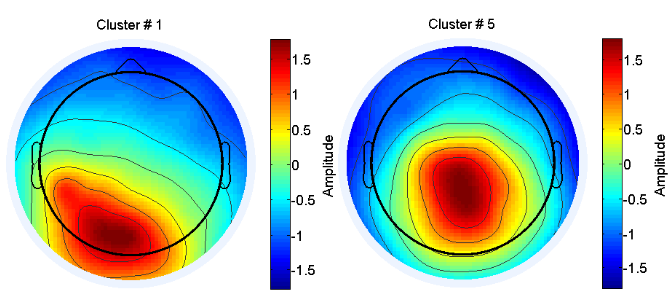

Fig. 8. Spatial maps of the clusters # 1 and # 5.

Table 1. # Contributors in the cluster #1 among 14 participants (from 1 to 14)

EEG	Music	Pulse clarity	Fluctuation entropy	Fluctuation centroid	Mode	Key
theta		N/A	2	2 5 12 14	3	N/A
alpha		7 10 13	13	2 13	1 3 10 14	5

Note: 'N/A' means that there are no contributors.

Table 2. # Contributors in the cluster #5 among 14 participants (from 1 to 14)

EEG	Music	Pulse clarity	Fluctuation entropy	Fluctuation centroid	Mode	Key
theta		4	N/A	3 10 13	4	10
alpha		14	8	8 9 10 12 14	6 13	12 13

Furthermore, in order to examine which oscillation appears in the central or occipital area along the scalp, the number of each contributor regarding each oscillation and each musical feature is reported for the clusters #1 and #5 in Tables-1 and 2. We find that the alpha and theta oscillations occur in the central and occipital area among most of participants, respectively. We also find that the musical feature, Fluctuation Centroid, is the most functional to elicit the ongoing EEG. Additionally, although Pulse Clarity and Fluctuation Entropy, Fluctuation Entropy & Fluctuation Centroid, and Mode and Key, are correlated ($p < 0.05$), the corresponding contributors in an interesting cluster do not entirely overlap in a pair of correlated features.

V. Conclusion and Discussion

In order to analyze ongoing EEG during the real-world experiences, this study formulates an approach with signal processing and analysis methods including digital filtering, ICA, brain imaging (analysis of residual dipole variance), clustering, acoustical feature extraction, spectral-temporal analysis, and correlation. Using this approach, we were able to provide the possibility to identify the brain regions involved in the processing of long-term acoustical features of modern tango from the ongoing EEG. Previously, the same was identified from fMRI of musicians [1], and naturalistic speech processing was localized from ongoing MEG [8]. As far as the authors are aware, this is the first complete formulation of an approach for the analysis of ongoing EEG in naturalistic and continuous music listening experiences.

ICA was used to decompose the ongoing EEG in this study. Therefore it is concerned the reliability of ICA components in such a new application. Here, the number of dipolar components is about 34% of all components when the threshold of the residual dipole variance is 15%. We checked the study of Delorme et al. [44] introducing the idea to validate the reliability of ICA components through the analysis of residual dipole variance. We found that such a percentage of the number of dipolar components extracted by InfomaxICA from ongoing EEG in this study is comparable to those of several ICA algorithms from single-trial ERP data in [44]. We think that although the ICA decomposition in this study is not perfect, it can be acceptable for the further analysis. In the future, we will learn more methods for artifacts rejection to test whether the ICA decomposition can be improved or not for the analysis of ongoing EEG, as well as testing higher sampling frequencies for the same purpose.

Indeed, there could be some time lags between music stimulus and the EEG [49, 50]. In this study, since we did not keep the reaction time in the experiment, the time lag is not available. We assume that the time lag cannot affect the time course of one EEG oscillation very much. The power of EEG oscillation is determined by every three-second EEG. Given the delay of dozens of or a few hundreds milliseconds to the three-second EEG, the power could not be affected significantly. In the future, it will be beneficial to measure the reaction time during the experiment. Then, we can precisely compare the effect of the delay of EEG for our proposed methodology.

In this study, the clustering was performed on the spatial maps. We found that the individual spatial maps in any cluster of interest were similar but their corresponding temporal courses of alpha or theta oscillations were not similar. This is validated by results in Tables 1 and 2 showing that the temporal courses of any EEG oscillation were significantly correlated with different musical features. Indeed, this is different from most of ERP studies where the temporal ICA components sharing similar spatial maps might be similar as well (this is indeed the basis for group ICA on ERPs [51]). We think the difference comes from the responses of participants during real-world experiences. EEG has very high temporal resolution. Such advantage can capture the different temporal evolutions of electrical responses of different people who are listening to the naturalistic and continuous music.

So far, there are very few publications reporting the analysis of ongoing EEG elicited by the naturalistic and continuous music. Nevertheless, the previous studies about ERPs elicited by the much shorter pieces of music [23-25] may be the suitable references for this study. Here, at the occipital area in Fig. 8 regarding the cluster #1, we found the alpha and theta oscillations relevant to the music stimulus. Such alpha activity has been denoted as the upper alpha which was elicited when listening to music [23] and such theta has not been reported so far. At the central area of the spatial map in Fig. 8 for the cluster #5, the alpha and theta oscillations associated with the music stimulus were discovered. Such theta activity has been reported in [23, 25] and such alpha is not discussed earlier. As the musical feature is a measure of the rhythmic complexity, our results are in line with previous findings that suggest heighted alpha activity due to increase in task demands [24], specifically while processing complex music [52, 53]. In another study of the ERPs elicited by short piano notes [54], the topographies of the ERPs revealed relevant brain responses in the central, posterior and occipital area along the scalp. Hence, the previous studies regarding EEG and music indicate that our findings from ongoing EEG elicited by naturalistic and continuous modern tango are plausible to be real brain activities, not technical artifacts.

ACKNOWLEDGMENT

This work was financially supported by TEKES (Finland) grant 40334/10 ‘Machine Learning for Future Music and Learning Technologies’. A.K. Nandi would like to thank TEKES for their award of the Finland Distinguished Professorship. Authors also thank anonymous reviewers for their important comments to improve the study.

REFERENCES

[1] 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.

[2] U. Hasson, Y. Nir, I. Levy, G. Fuhrmann and R. Malach, "Intersubject synchronization of cortical activity during natural vision," Science, vol. 303, pp. 1634-1640, 2004.

[3] J. D. Haynes and G. Rees, "Decoding mental states from brain activity in humans," Nat. Rev. Neurosci., vol. 7, pp. 523-534, 2006.

[4] J. P. Kauppi, I. P. Jaaskelainen, M. Sams and J. Tohka, "Inter-subject correlation of brain hemodynamic responses during watching a movie: localization in space and frequency," Front. Neuroinform, vol. 4, 2010.

[5] K. N. Kay, T. Naselaris, R. J. Prenger and J. L. Gallant, "Identifying natural images from human brain activity," Nature, vol. 452, pp. 352-355, 2008.

[6] H. J. Spiers and E. A. Maguire, "Decoding human brain activity during real-world experiences," Trends Cogn. Sci., vol. 11, pp. 356-365, 2007.

[7] A. Hyvarinen, P. Ramkumar, L. Parkkonen and R. Hari, "Independent component analysis of short-time Fourier transforms for spontaneous EEG/MEG analysis," Neuroimage, vol. 49, pp. 257-271, 2010.

[8] M. Koskinen, J. Viinikanoja, M. Kurimo, A. Klami, S. Kaski and R. Hari, "Identifying fragments of natural speech from the listener's MEG signals," Hum. Brain Mapp., 2012.

[9] H. Pan, J. Epstein, D. A. Silbersweig and E. Stern, "New and emerging imaging techniques for mapping brain circuitry," Brain Res. Rev., vol. 67, pp. 226-251, 2011.

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

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

[12] 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.

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

[14] 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.

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

[16] S. Debener, F. Minow, R. Emkes, K. Gandras and M. de Vos, "How about taking a low-cost, small, and wireless EEG for a walk?" Psychophysiology, vol. 49, pp. 1617-1621, 2012.

[17] 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, 2009.

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

[19] 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.

[20] 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, 2001.

[21] 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, 2004.

[22] G. Buzsáki, Rhythms of the Brain. Oxford University Press, 2006.

[23] D. Sammler, M. Grigutsch, T. Fritz and S. Koelsch, "Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music," Psychophysiology, vol. 44, pp. 293-304, 2007.

[24] 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, 2011.
[25] A. J. Shahin, L. J. Trainor, L. E. Roberts, K. C. Backer and L. M. Miller, "Development of auditory phase-locked activity for music sounds," J. Neurophysiol., vol. 103, pp. 218-229, 2010.
[26] S. Makeig, T. P. Jung, A. J. Bell, D. Ghahremani and T. J. Sejnowski, "Blind separation of auditory event-related brain responses into independent components," Proc. Natl. Acad. Sci. U. S. A., vol. 94, pp. 10979-10984, 1997.
[27] S. Makeig, M. Westerfield, T. P. Jung, J. Covington, J. Townsend, T. J. Sejnowski and E. Courchesne, "Functionally independent components of the late positive event-related potential during visual spatial attention," J. Neurosci., vol. 19, pp. 2665-2680, 1999.
[28] F. Cong, Y. Huang, I. Kalyakin, H. Li, T. Huttunen-Scott, H. Lyytinen and T. Ristaniemi, "Frequency Response based Wavelet Decomposition to Extract Children's Mismatch Negativity Elicited by Uninterrupted Sound," J. Med. Biol. Eng., vol. 32, pp. 205-214, 2012.
[29] F. Cong, P. H. Leppanen, P. Astikainen, J. Hämäläinen, J. K. Hietanen and T. Ristaniemi, "Dimension reduction: Additional benefit of an optimal filter for independent component analysis to extract event-related potentials," J. Neurosci. Methods, vol. 201, pp. 269-280, 2011.
[30] R. Vigario and E. Oja, "BSS and ICA in Neuroinformatics: From Current Practices to Open Challenges," IEEE Reviews in Biomedical Engineering, vol. 1, pp. 50-61, 2008.
[31] 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.
[32] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis," J. Neurosci. Methods, vol. 158, pp. 300-312, 2006.
[33] O. Lartillot and P. Toiviainen , "MIR in matlab (II): A toolbox for musical feature extraction from audio," in Proceeding of International Conference on Music Information Retrieval 2007, 2007, pp. 237-244.
[34] F. Cong, I. Kalyakin and T. Ristaniemi, "Can back-projection fully resolve polarity indeterminacy of ICA in study of ERP?" Biomed. Signal Process. Control, vol. 6, pp. 422-426, 2011.
[35] F. Cong, I. Kalyakin, C. Zheng and T. Ristaniemi, "Analysis on Subtracting Projection of Extracted Independent Components from EEG Recordings," Biomed. Tech., vol. 56, pp. 223-234, 2011.
[36] A. Hyvärinen, J. Karhunen and E. Oja, Independent Component Analysis. New York: John Wile & Sons Inc., 2001.
[37] J. Onton, M. Westerfield, J. Townsend and S. Makeig, "Imaging human EEG dynamics using independent component analysis," Neurosci. Biobehav. Rev., vol. 30, pp. 808-822, 2006.
[38] M. De Vos, J. D. Thorne, G. Yovel and S. Debener, "Let's face it, from trial to trial: Comparing procedures for N170 single-trial estimation," Neuroimage, vol. 63, pp. 1196-1202, 2012.
[39] Y. Wang, Y. T. Wang and T. P. Jung, "Translation of EEG spatial filters from resting to motor imagery using independent component analysis," PLoS One, vol. 7, pp. e37665, 2012.
[40] A. Cichocki and S. Amari, Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. John Wile & Sons Inc.: Chichester, 2003.
[41] F. Cong, I. Kalyakin, T. Huttunen-Scott, H. Li, H. Lyytinen and T. Ristaniemi, "Single-trial based independent component analysis on mismatch negativity in children," Int. J. Neural Syst., vol. 20, pp. 279-292, 2010.
[42] J. Himberg, A. Hyvärinen and F. Esposito, "Validating the independent components of neuroimaging time series via clustering and visualization," Neuroimage, vol. 22, pp. 1214-1222, 2004.
[43] A. J. Bell and T. J. Sejnowski, "An information maximisation approach to blind separation and blind deconvolution," Neural Computation, vol. 7, pp. 1159, 1995.
[44] A. Delorme, J. Palmer, J. Onton, R. Oostenveld and S. Makeig, "Independent EEG sources are dipolar," PLoS One, vol. 7, pp. e30135, 2012. [45] C. M. Bishop, Pattern Recognition and Machine Learning. Singapore: Springer, 2006.
[46] L. Kaufman and P. Rousseeuw, Finding Groups in Data: An Introduction to Cluster Analysis. San Francisco: Wiley, 2005.
[47] J. Rissanen, "Modeling by the shortest data description," Automatica, vol. 14, pp. 465-471, 1978.
[48] D. M. Groppe, T. P. Urbach and M. Kutas, "Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review," Psychophysiology, vol. 48, pp. 1711-1725, 2011.
[49] 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, Apr 9, 2008.
[50] S. J. Aiken and T. W. Picton, "Envelope and spectral frequency-following responses to vowel sounds," Hear. Res., vol. 245, pp. 35-47, 2008.
[51] F. Cong, Z. He, J. Hämäläinen, P. H. T. Leppänen, H. Lyytinen, A. Cichocki and T. Ristaniemi, "Validating Rationale of Group-level Component Analysis based on Estimating Number of Sources in EEG through Model Order Selection," J. Neurosci. Methods, vol. 212, pp. 165-172, 2013.
[52] W. Gunther, R. Steinberg, P. Streck, J. P. Banquet, I. Bscheid, L. Raith, R. Riedel, U. Klages and I. Stiltz, "Brain dysfunction in psychiatric patients during music perception measured by EEG mapping: relation to motor dysfunction and influence of neuroleptic drugs," Eur. Neuropsychopharmacol., vol. 1, pp. 143-155, 1991.
[53] M. H. Ruiz, S. Koelsch and J. Bhattacharya, "Decrease in early right alpha band phase synchronization and late gamma band oscillations in processing syntax in music," Hum. Brain Mapp., vol. 30, pp. 1207-1225, 2009.
[54] S. Elmer, M. Meyer and L. Jancke, "The spatiotemporal characteristics of elementary audiovisual speech and music processing in musically untrained subjects," Int. J. Psychophysiol., 2011.