This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Mahini, Reza; Li, Yansong; Ding, Weiyan; Fu, Rao; Ristaniemi, Tapani; Nandi, Asoke K.; Chen, Guoliang; Cong, Fengyu

Title: Determination of the Time Window of Event-Related Potential Using Multiple-Set Consensus Clustering

Year: 2020

Version: Published version

Copyright: © 2020 Mahini, Li, Ding, Fu, Ristaniemi, Nandi, Chen and Cong

Rights: CC BY 4.0

Rights url: https://creativecommons.org/licenses/by/4.0/

Please cite the original version:
Mahini, R., Li, Y., Ding, W., Fu, R., Ristaniemi, T., Nandi, A. K., Chen, G., & Cong, F. (2020). Determination of the Time Window of Event-Related Potential Using Multiple-Set Consensus Clustering. Frontiers in Neuroscience, 14, Article 521595. https://doi.org/10.3389/fnins.2020.521595
Determination of the Time Window of Event-Related Potential Using Multiple-Set Consensus Clustering

Reza Mahini1,2, Yansong Li3,4, Weiyian Ding5, Rao Fu1, Tapani Ristaniemi6†, Asoke K. Nandi6, Guoliang Chen5* and Fengyu Cong1,2,7,8*

1 School of Biomedical Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China, 2 Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland, 3 Reward, Competition and Social Neuroscience Lab, Department of Psychology, School of Social and Behavioral Sciences, Nanjing University, Nanjing, China, 4 Institute for Brain Sciences, Nanjing University, Nanjing, China, 5 Department of Psychiatry, Chinese PLA 967th Hospital, Dalian, China, 6 Department of Electronic and Computer Engineering, Brunel University London, Uxbridge, United Kingdom, 7 School of Artificial Intelligence, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China, 8 Key Laboratory of Integrated Circuit and Biomedical Electronic System, Liaoning Province, Dalian University of Technology, Dalian, China

Clustering is a promising tool for grouping the sequence of similar time-points aimed to identify the attention blocks in spatiotemporal event-related potentials (ERPs) analysis. It is most likely to elicit the appropriate time window for ERP of interest if a suitable clustering method is applied to spatiotemporal ERP. However, how to reliably estimate a proper time window from entire individual subjects’ data is still challenging. In this study, we developed a novel multiset consensus clustering method in which several clustering results of multiple subjects were combined to retrieve the best fitted clustering for all the subjects within a group. Then, the obtained clustering was processed by a newly proposed time-window detection method to determine the most suitable time window for identifying the ERP of interest in each condition/group. Applying the proposed method to the simulated ERP data and real data indicated that the brain responses from the individual subjects can be collected to determine a reliable time window for different conditions/groups. Our results revealed more precise time windows to identify N2 and P3 components in the simulated data compared to the state-of-the-art methods. Additionally, our proposed method achieved more robust performance and outperformed statistical analysis results in the real data for N300 and prospective positivity components. To conclude, the proposed method successfully estimates the time window for ERP of interest by processing the individual data, offering new venues for spatiotemporal ERP processing.

Keywords: multi-set consensus clustering, time window, event-related potentials, microstates analysis, cognitive neuroscience

INTRODUCTION

The event-related potentials (ERPs) carry important information about the cognitive process evoked by the brain response in milliseconds of the temporal domain. Almost all the ERP components are influenced by the attention corresponding to the latencies from the individual and a group of subjects (Luck and Kappenman, 2012). The latencies of ERP components can be
considered as a stable brain electric field configuration (topography map) in milliseconds associated with the specific psychological process (i.e., attention module) (Lehmann, 1990). Moreover, measuring the ERP of interest undertakes a fundamental role in identifying and interpreting the cognitive process in the experiment. The most common approach to measure the magnitude and timing of the ERP of interest is to investigate the amplitude and the latency of peak voltage in the experimentally defined time window. Thereby, an important issue in the analysis of ERPs is how to define or select time windows. This influences both identifying components and performing statistical analyses. Hence, if the time window is not appropriately defined, the comparison between different conditions/groups can lead to unreliable and wrong psychological interpretations (Luck and Gaspelin, 2017).

The traditional ERP approach is to obtain the mean of measured potentials over a fixed and/or experimenter defined time window. The assumption is that the brain electric field configuration is stable for different conditions/groups, although this assumption is not empirically verified. Apart from widely used conventional ERP techniques such as latency peak and mean amplitude, numerous studies have used moving time-window technique and high-resolution time-bin analysis (e.g., each 5 ms) for measuring the peak (Van Overwalle et al., 2009; Mu and Han, 2010; Wills et al., 2014). Although moving time-window or point-by-point analysis in spatiotemporal ERP can provide more fine-grained temporal characterization and significant statistical results (Rothstein et al., 2010), they can dramatically increase the probability of reporting errors (Luck and Gaspelin, 2017). In the above reviewed methods, the variety of responses, which dynamically influence the duration of time windows in different conditions/groups, are neglected.

Another group of researchers investigated the brain response states by analyzing the topographical changes (Lehmann, 1989, 1990; Lehmann et al., 1994; Micah et al., 2009) to determine the components of interest. The underlying assumption is that the electric field configuration does not change randomly as a function of time, despite exhibiting stability for tens to hundreds of milliseconds involving intervals of topographic instability (Lehmann et al., 1987; Murray et al., 2008). The clustering of spatiotemporal electroencephalogram (EEG)/ERP was used to capture template maps (i.e., topographies found by the clustering) which identifies the recorded signal (Lehmann, 1989, 1990). Hypothetically, the brain state (i.e., the brain electric field configuration) does not change during a specific response time (Lehmann, 1990; Pascual-Marqui et al., 1995; Lehmann et al., 2009). Consequently, the spatial correlation of corresponding topographies of the time-points in the cluster map is close to 1 (Pourtois et al., 2008). Two clustering algorithms in EEG/ERP research, namely, modified k-means (Pascual-Marqui et al., 1995) and agglomerate hierarchical clustering (AAHC; Tibshirani and Walther, 2005; Murray et al., 2008) were predominantly used in EEG/ERP researches. Two global measurements together, namely, global field power (GFP) and the global map dissimilarity (GMD), and the global explained variance (GEV) of the template maps (the most important cluster maps), for quantifying the template maps, were applied. Furthermore, the topographical analysis for spatiotemporal ERPs using clustering methods has been explored in several studies (Murray et al., 2008; Micah et al., 2009; Koenig et al., 2014). So far in the aforementioned microstates analysis studies (Michel and Koenig, 2018), determination of template cluster maps with higher explained variance and post hoc determination of microstates by fitting those maps to the data (topography maps) were used. As a result, the time-points are clustered based on their similarity in the electrode field configuration. Alternative methods, for cluster or factor analysis, such as optimized k-means with genetic algorithm and principal component analysis (PCA) (Williams et al., 2015), topographic pattern analysis, and PCA in high-density ERP (Pourtois et al., 2008) were utilized to determine the most dominant spatial components from the map series. Although independent component analysis and PCA are standard methods and are used for decomposition of the EEG/ERP with cluster analysis, the determination of the event of the interest is subjective instead of being the objective exploration of ERP.

Importantly, finding the suitable time window for measuring the ERP of interest using microstates analysis has also been studied in the numerous literature (Tzovara et al., 2012; Cacioppo et al., 2014; Koenig et al., 2014; Khanna et al., 2015; Mahe et al., 2015). The time window has been determined by testing point-time by point-time, the topographical ANOVA analysis, and microstate classes on momentary grand-mean maps (Koenig et al., 2011). Some recent studies, for example, have explored the most suitable time window from the most fitted microstate maps via the clustering of spatiotemporal ERP by comparing the ERPs of individual subjects with the obtained ERPs of clustering from grand average data (Bailey et al., 2019; Berchio et al., 2019; Ruggeri et al., 2019). Although obtaining global optimal cluster maps by clustering both group (grand average) and individual datasets assigning time-points to template maps is a straightforward solution (Michel and Koenig, 2018), it is challenging to set of template maps from grand average ERP, which reliably represent individual subjects brain responses.

Consensus clustering, as a reliable and stable clustering method, has been successfully used for processing biological data (Monti et al., 2003; Abu-Jamous et al., 2013, 2015a; Liu et al., 2015; Mahini et al., 2017), human brain functional magnetic resonance imaging and EEG data processing (Liu et al., 2017a, b; Song et al., 2019), and multidataset consensus clustering (Filkov and Skiena, 2004; Hoshida et al., 2007; Abu-Jamous et al., 2015b; Liu et al., 2015). However, there has been little discussion about the role of multidataset consensus clustering on individual data from spatiotemporal ERP aimed to identify the ERP components. This is critical because of the difference between the subjects regarding the response time and delay and difference in the quality of recorded data. Therefore, a robust method is required for processing information about the subjects.

The rationale of the current study is to investigate three major points; first, in the ERP experiment, several ERP components are inevitably generated; however, a few of them are targeted, which are more probably elicited if the ERP experiment is run again (Kappenman and Luck, 2012b). Those targeted ERP components are more probably elicited among multiple subjects.
The proposed method isolates reliable time windows for ERP of interest for each condition/group. Second, essentially, even after the well-done preprocessing of the collected data, there are still some remaining interferences and some overlapped brain activity with the ERP of interest in the time domain. Therefore, it is practically expected the time window for measuring the amplitude of the ERP of interest includes information of the ERP. One strategy is to check whether the consecutive multiple topographies of time-points are similar or not. If they are similar enough, they come from the same brain activity of the ERP in terms of the linear transformation model of EEG. Thus, such a time window should be determined. Since the time window contains mostly the ERP of interest, the analysis of the brain response can be more accurate. This can result in a better understanding of cognitive processes. Finally, the ERP signal is elicited from numerous similar responses from the subjects. Defining the ERP of interest from the clustering of grand average data neglects the information about individual subjects. Thereby, the new methodology explores the ERP of interest from individual subjects using a multisubject consensus clustering.

In this article, we develop a stabilized multiple-subject consensus clustering (from the multiset consensus clustering family) approach for reliably clustering spatiotemporal ERP data in both individual subjects and group levels. This can provide a novel mechanism to explore the cognitive functions in ERP/EEG data. Furthermore, we use a newly proposed time-window determination method to obtain the most suitable time window for a given ERP of interest. We do expect the new methodology can retrieve the consistent response among the subjects in a group to discover a reliable time window for the ERP of interest. To assess the efficiency and reliability of our method, the proposed method is applied to simulated and the prospective memory experiment data (Chen et al., 2015). The proposed method has been tested to identify two state-of-the-art ERPs, namely, N2 and P3 components in simulated data, and isolating N300 and prospective positivity components in the real data.

**MATERIALS AND METHODS**

This section describes first two ERP datasets including conducted simulated data and real data. Then, our proposed method is described in detail. Finally, two classes of statistical analysis for assessing the studied methods are explained.

**ERP Studies**

**Simulated ERP Data**

We conducted a simulated ERP data using the BESA dipole simulator\(^1\) for assessing the performance of the studied clustering methods aimed to identify the predefined ERP components. Entirely, six components (i.e., P1, N1, P2, N2, P3, and N4) and two conditions (i.e., “Cond1” and “Cond2”) from a group of 20 subjects were defined. A simulated scalp with 65 electrodes was used for representing the spatial (i.e., topographic) information. Each trial was epoched from 100-ms prestimulus to 600-ms poststimulus at a sampling rate of 429 Hz. The averaged reference method was used for referencing. The topography maps of the components and corresponding waveforms are

\(^1\)https://www.besa.de/products/besa-simulator/besa-simulator-overview/
shown in Figure 1. Among the defined components, we studied N2 referring to the maximum negative voltage in 201- to 265-ms poststimulus [i.e., it was defined in 175–292 ms via simulator (Figure 1)]. The time window was calculated using the signed area measurement method (Sawaki et al., 2012). Similarly, P3 component refers to the positive response (266–357 ms) poststimulus (i.e., defined in 240–385 ms according to Figure 1). Meanwhile, the signal was manipulated using the MATLAB function awgn (i.e., adding white Gaussian noise) to add a reasonable noise (i.e., signal-to-noise ratio = 20 dB) on signal power measured for each simulated dataset as a whole. Furthermore, random movement of two ERPs (e.g., changing the original signal by randomly increasing/decreasing maximum five time-points) was applied to the original signal from the 20 individuals’ data. The electrode sites for measuring statistical amplitude power differences were defined as P6/PO4 and CPz/Cz for N2 and P3, respectively.

Real ERP Data

The prospective memory experiment (Chen et al., 2015) data were used as real ERP data to assess the performance of the proposed method. Following the prior study, the experiment data included 20 symptomatically remitted patients, i.e., with schizophrenia (RS) and 20 healthy control (HC) participants. Two tasks, namely, prospective memory (PM) and ongoing task, were investigated. The EEG data were recorded with 32 electrodes (SynAmps amplifier, NeuroScan) and epoched from 200-ms prestimulus to 1,000-ms poststimulus. Furthermore, a 30 Hz (24 dB/octave) digital low-pass filter was applied. Two target ERP components, N300 and prospective positivity components, were studied. The N300 referred to the maximum negative voltage, over the occipital region, hypothetically between 190 and 400 ms, and the prospective positivity represented the maximum positive voltage, over the parietal region, and between 400 and 1,000 ms.

Proposed Method

The graphical explanation of the proposed method is illustrated in Figure 2. Besides, Procedure 1 and Procedure 2 are presented for a better representation of the new methodology. Noteworthy to mention that we have employed a mechanism to obtain the optimal number of clusters by, first, running the consensus clustering many times followed by determining the optimal number of clusters based on the quality of obtaining time windows (Mahini et al., 2019). The details of the proposed method are given as follows:

Procedure 1: Proposed Method
Inputs: ERP data, ERPs of interest (experimental intervals)
Outputs: Time windows
Procedure {
    Step 1: Temporal concatenating datasets for each individual subject;
    Step 2: Stabilization and generation;
    Step 3: Multilevel consensus clustering
        Individual level consensus clustering;
        Group level consensus clustering;
    FOR each ERP of interest
    Step 4: Time-window determination;
} End of FOR

Dataset for Clustering

The collected multiple data points by a high-dense EEG sensor array consist of the spatial topographies of brain activities (i.e., each time-point corresponds to a topography). We have investigated the spatiotemporal ERP data where the time-points are clustered based on their topographical similarity. For each subject, a larger dataset was yielded from temporal concatenating (Murray et al., 2008; Calhoun et al., 2009) the associated datasets from all conditions together. For example, given a subject’s ERP data from 300 time-points, 2 conditions, and 65 electrodes, the temporal concatenated dataset with a dimension of 600 × 65 is used for clustering. Therefore, the samples for clustering individual data are the time-points, and the features are represented by the topography (i.e., the electrode field configuration). The goal of clustering is to find the consecutive time-points sharing similar topographies in which the neural responses remain stable for periods of time called time window.

Stabilization and Generation

We utilized the cluster-based similarity partitioning algorithm method (Karypis and Kumar, 1998; Nguyen and Caruana, 2007) as the consensus function based on pairwise similarity measurement between partitions. This function was used for each level of consensus clustering and the stabilization step of the proposed method. Before the generation step, two important issues, consensus clustering configuration and stabilized generation, are necessary to be investigated. Several clustering methods were considered for selecting the appropriate configuration of consensus clustering. Hence, k-means (Pascual-Marqui et al., 1995; Pena et al., 1999) and hierarchical clustering (Tibshirani and Walther, 2005) with correlation similarity function, fuzzy c-means (FCM; Bezdek, 1981), self-organizing maps (SOMs; Kohonen, 1990), diffusion map spectral clustering (Sipola et al., 2013) consisting k-means with Euclidean similarity, and modified k-means (Pascual-Marqui et al., 1995), and AAHC (Murray et al., 2008) using spatial correlation, were used for the generation purpose. Thereby, for appropriate consensus clustering configuration, modified k-means was used as a benchmark [i.e., the accepted clustering method in many studies (Michel and Koenig, 2018)] to be compared with other studied clustering methods. The clustering methods with higher mutual similarities with modified k-means in the majority of clustering results of individuals data (e.g., ≥50% of the subjects), were selected using in the generation phase. Rand index (Strehl and Ghosh, 2003; Meila, 2007) was used to measure the mutual similarity between the results of each clustering method on individual data and modified k-means. Rand index can be calculated using the following equation:

\[
\mathcal{R}(L, L') = \frac{N_{11} + N_{00}}{n(n-1)/2}
\]
where \( n \) denotes the number of observations and \( N_{ij0} \) denotes the number of object pairs in different clusters from both \( L \) and \( L' \) clusterings. While \( N_{ij1} \) denotes the number of object pairs in the same clusters in \( L \) and \( L' \).

Additionally, a stabilization procedure based on consensus clustering was designed for the clustering generation of consensus clustering (at the subject level). The stable clustering refers to the clustering results in which the mutual similarity between two or more clustering results is closed to 1 in theory. To measure stability, a mechanism based on the testing similarity of two clustering results was utilized. If they are highly similar, the clustering method is robust. The consensus clustering of grand average ERP data from multiple runs of each stochastic clustering method (e.g., from 2 to 20 repeats that can be changed if necessary) was employed to find the appropriate number of repetitions to get stable clustering. The optimal number of repetitions should satisfy the following two conditions:

\[
\max(|R_r - R_{r-1}|, |R_r - R_{r+1}|) \leq \varepsilon \tag{2}
\]

where

\[
R_r = \mathcal{R}\left(L^{s-r}, L^{s-(r-1)}\right)
\]

\[
L^{s-r} = \arg \max_{L \in \mathcal{R}} \sum_{r=2}^{Mr} \Gamma(L_r)
\]

and

\[
\min(R_{r-1}, R_r, R_{r+1}) \geq \tau \tag{3}
\]

where \( \Gamma \) denotes the consensus function, \( L^{s-r} \) denotes the consensus clustering results from \( r \) repetitive results (i.e., maximum repeats denotes by \( Mr \)) of stochastic clustering method, which is indicated by \( L_r \), and \( \bar{X} \) denotes the grand average from the individual datasets. Furthermore, \( R_r \) denotes the mutual similarity between the consensus clustering results from \( r \) and \( r - 1 \) repetitions. Thus, a proper number of repetitions is determined by measuring the mutual similarity among the results of consensus clustering. In other words, the optimal repetition option is selected when the mutual similarity between \( r - 1 \) and \( r \), and between \( r \) and \( r + 1 \) reaches a suitable similarity threshold (e.g., \( \tau \geq 90 \)), and the change among mutual similarities tends to very small values (e.g., \( \varepsilon \leq 0.03 \)).

**Multilevel Consensus Clustering**

A two-level consensus clustering was utilized for finding the best fitted clustering from individual subjects. The proposed two-level multisubject consensus clustering is explained by the following notations:

Let \( S = \{S_1, S_2, \ldots, S_p\} \) denotes a set of subjects from a group, and \( X = \{x_1, x_2, \ldots, x_n\} \) denotes a set of time-points for individual data, in which each time-point \( x_s = \{e_1, e_2, \ldots, e_f\} \), \( s = 1, 2, \ldots, n \) (\( f \) denotes the number of electrodes) is a vector of features/channels (i.e., it can be represented in the spatial dimension as a topography map).

Besides, \( L_i^j = \{C_{1,i}^j, C_{2,i}^j, \ldots, C_{k,i}^j\} \) represents the clustering results for jth clustering method \( j = 1, 2, \ldots, m \), for ith subject, \( i = 1, 2, \ldots, p \) with \( k \) number of clusters. Thus, \( C_{w,i}^j \) is defined as with cluster, \( w = 1, 2, \ldots, k \) from jth method for ith subject. The
result of the first-level clustering for each of individual datasets is denoted as:

\[ L_i^{s-opt} = \arg \max_{L \in \mathcal{X}} \sum_{j = 1}^{m} \Gamma \left( t_i^j \right) \]  

(4)

where, \( L_i^{s-opt} \) denotes the consensus clustering results of \( i \)th subject from all possible \( k \)-partitions on \( X \). At the second level, another consensus clustering is used on the first level clustering results across the subjects (i.e., in the group level), which is defined as:

\[ L^{s*\cdot opt} = \arg \max_{L \in \mathcal{X}} \sum_{j = 1}^{p} \sum_{j = 1}^{m} \Gamma \left( L_i^{s-opt} \right) \]  

(5)

where, \( L^{s\cdot opt} \) denotes the result of consensus clustering across the subjects.

Taken as a whole, the optimal ensemble clustering across the subjects can be noted by:

\[ L^{s*\cdot opt} = \arg \max_{L \in \mathcal{X}} \sum_{j = 1}^{p} \sum_{j = 1}^{m} \Gamma \left( L_i^{s-opt} \right) \]  

(6)

To provide a better sense of implementation of the proposed method, the multisubject consensus clustering was implemented in MATLAB platform, as demonstrated in Figure 2 and Procedure 1.

**Time Window Determination**

The time window determination procedure explores the measurement time window by analyzing the temporal and spatial characteristics of the result cluster maps. The inner-similarity of the candidate cluster map (the maps in the experimental measurement area) and their overlapping with the defined experimental time interval, were considered to estimate the proper time windows. First, the inner-similarity of candidate maps is calculated aimed to detect those with the consecutive time-points with a high spatial correlation. The inner-similarity of a cluster map is the mean of correlation coefficients between topography maps of each two different time-points. More in detail, to calculate the inner-similarity of a cluster map, first, the spatial correlation coefficient (Murray et al., 2008; Micah et al., 2009) of time-points was calculated. Therefore \( Cor_{r,u} \) denotes the correlation coefficient between the topographical maps of \( u \) and \( v \) as two time-points in the cluster map. Then, for each row, the distance matrix can be calculated as:

\[ D_r = d \left( Cor_{r,u}, Cor_{r,v} \right), \quad u \neq v \]  

(7)

where, \( D \) denotes the distance matrix in which each row is calculated by the distance between each element in the row and \( Cor_{r,v} \) (i.e., self-correlation) in correlation matrix (Cor). To variance-stabilizing transformation of the calculated correlation, fisher z-transform (Fisher, 1921) was used for each vector \( D_r \) (i.e., every row of distance matrix) before calculating the mean of the distance matrix \( D_{avg} \). Finally, an inverse z-transform of \( D_{avg} \) was used for calculating inner-similarity as shown below:

\[ InnSim = 1 - D_{avg} \]  

(8)

Hypothetically, in the ERP component, the spatial correlation between the time-points is close to 1 indicating consecutive time-points that represent a cognitive process. Therefore, among the candidate cluster maps, the cluster maps with higher inner similarity than the threshold (e.g., \( \geq 0.90 \)) were selected for overlap testing. We have selected a realistic choice of 0.9 as a satisfactory threshold for time-window qualification. Next, among those cluster maps, the cluster map with the greatest inner-similarity and overlapping was selected as the best suitable cluster map for representing the time window (i.e., via the properties (start, end, and duration)). More details for implementing the time-window selection method are presented in Procedure 2.

**Procedure 2: Time-Window Determination**

**Input:** Clustering result, ERPs of interest (experimental intervals)

**Output:** Time windows

**Procedure**

1. **Step 1.** Detecting the candidate cluster maps;
   
   **FOR** each candidate map
   
   **Step 2.** Calculating inner-similarity and overlapping;
   
   **Step 3.** Detecting cluster maps with high inner-similarity;
   
   **Step 4.** Selecting higher overlapping within maps;
   
   **End of FOR**

**Statistical Analysis**

Two classes of \( p \) values based statistical measurements were used to evaluate the performance of the proposed method. First, two one-sided tests (TOST; Rogers et al., 1993; Harms and Lakens, 2018) was performed on simulated data to test the similarity between ground truth and estimated time windows by measuring the obtained time-window properties (start, end, and duration). Second, a statistical power analysis was used by employing repeated measures ANOVA for both simulated and real data. Further, for testing the robustness of those methods, the statistical analysis results were calculated on over 50 independent runs of the studied methods. Overall, we tried to assess the meaningfulness, accuracy, and robustness of the proposed methodology.

The TOST test was accomplished by setting equivalence margin \([-\delta, \delta]\) in \([-5, 5]\) ms (can vary depending on the dataset and quality of discriminability). Two composite null hypotheses tested the assumption of the differences: \( H_0_1: (\mu_1 - \mu_2) \leq -\delta \) and \( H_0_2: (\mu_1 - \mu_2) \geq \delta \), where \( \mu_1, \mu_2 \) are the mean of each series in the comparison (e.g., the estimated start points from all the individual subjects in a group and corresponding ground truth start points). When both null hypotheses can be statistically rejected, it can be concluded that the observed effect falls within the equivalence margins and practically equivalent (Seaman and Serlin, 1998). In other words, the difference between the mean of the estimated values and the corresponding ground truth values should not exceed the equivalence margins. Furthermore, a repeated-measures ANOVA for the simulated data with the within-subject factor: task (“Cond1” and “Cond2”) was considered for statistically analyzing N2 component in the
electrode sites: P6/PO4 for N2 and CPz/Cz for P3. The test was applied to the mean amplitude of N2 and P3 in the estimated time windows separately. Similarly, the statistical power analysis for real data was carried out via repeated measures ANOVA (i.e., mixed 2 × 2) with the addition of a between-subject factor: group (RS and HC) and the within-subject factor: task (PM and ongoing). The test was applied to the mean amplitude of N300 and prospective positivity. The selection of electrodes was based on prior ERP findings (Chen et al., 2015). Specifically, the amplitude of N300 over the occipital region (electrodes: O1/Oz/O2) and prospective positivity over the parietal region (electrodes: P3/Pz/P4) were measured. Statistical comparisons were made at p values of p < 0.05 for both data.

RESULTS

To achieve the appropriate clustering result, several important parameters were adjusted, (i) determination of the optimal number of clusters: following our previous study (Mahini et al., 2019), the appropriate number of clusters for simulated and real data was determined in five and six cluster maps, respectively. (ii) The configuration of the proposed consensus clustering among the studied clustering methods (addressed in “Stabilization and Generation”), k-means, hierarchical clustering, AAHC, and modified k-means methods were applied to the simulated data. Similarly, k-means, FCM, SOMs, diffusion map spectral clustering, AAHC, and modified k-means methods were selected for the clustering of real data (Table 1). (iii) Generating stabilized clustering from stochastic clustering methods: following (section “Stabilization and Generation”) the optimal repeat for modified k-means and standard k-means was obtained in five and seven repeats for the simulated data (Figure 3). Likewise, those clustering methods met stability in seven repetitions in real data. Furthermore, a realistic inner-similarity threshold (e.g., ≥0.90) and a sufficient number of time-points for selecting the candidate cluster maps, e.g., a minimum of 60 to 100 ms (Grieder et al., 2016; Koenig and Brandeis, 2016) were determined.

Results of Simulated ERP Data

We applied the proposed consensus clustering in the simulated data aimed to illustrate all the predefined ERP components. The clustering in seven cluster maps successfully isolated all predefined six components (Figure 4) P1, N1, P2, N2, P3, and N4 correspond with the cluster maps 3, 5, 6, 1, 7, and 2, respectively. Note that cluster map 4 refers to the brain state before stimulus onset and does not present any predefined ERP component.

Time Windows and Topographies for ERPs of Interest

Figure 5 illustrates the clustering results and the elicited N2 and P3 components (from one random execution), including the corresponded topography maps and the spatial correlation of time-points obtained by the proposed method on the simulated data. Figure 5A indicates that the N2 component in Cond1 and Cond2 are elicited by cluster maps 5 (marked blue). Likewise, Figure 5B illustrates that the P3 component is identified by the microstate map 1 (marked orange) in both conditions. These results reveal that a significant main effect of task (p < 0.0001) was detected in N2 in the duration of microstate maps. Similarly, a significant main effect of task (p < 0.0001) was detected in the P3 component. For both components, the measured amplitudes were greater in Cond2. This reveals that the N2 and P3 components seem to be distinctly elicited by the proposed method in the simulated data.

Comparison Between Estimated and Ground Truth Time Windows

The proposed method was compared with the state-of-the-art clustering methods, namely, modified k-means and AAHC, in spatiotemporal ERP clustering. Our time-window selection method was applied to the clustering results (i.e., proposed consensus clustering, modified k-means, and AAHC results) to identify each ERP of interest. The Start, End, and Duration

| Data          | Group | KMS | HC | FCM | SOM | DSC | AAHC |
|---------------|-------|-----|----|-----|-----|-----|------|
| Simulated data| G1    | 19  | 14 | 0   | 0   | 0   | 20   |
| Real data     | RS    | 19  | 9  | 17  | 17  | 15  | 20   |
|               | HC    | 19  | 11 | 19  | 19  | 15  | 18   |

The marked methods with bold font are selected where they achieved higher similarity (rand index) for the majority of individual data (e.g., ≥50% of subjects and similarity ≥ 0.7). RS, remitted schizophrenia; HC, healthy control; G1, simulated group; KMS, k-means; HC, hierarchical clustering; FCM, fuzzy c-means; SOM, self-organizing map; DSC, diffusion maps spectral clustering; and AAHC, atomize and agglomerate hierarchical clustering.

FIGURE 3 | The illustration of the clustering method selection by calculating the similarity of the results with the modified k-means method for individual data.

FIGURE 4 | The illustration of the stability test in 20 runs of studied clustering methods to the grand average ERP data. Dash lines demonstrate the original clustering method stability and the continuous lines illustrate the corresponding stabilized version for each studied clustering method behavior for the range of repetition (e.g., from 2 to 20). (A) Stabilizing in the simulated data and (B) stabilizing in the real data.
parameters of estimated time windows were compared with that of the ground truth time windows (obtained from the simulation) on the clustering results of individual data for testing the accuracy. The TOST result (Table 2) for N2 component from clustering methods illustrates that the null hypothesis was rejected for the proposed method, modified $k$-means, and AAHC for all parameters in both conditions except End in Cond1 for AAHC. Similarly, the null hypothesis was rejected in all parameters except Duration in both conditions for the proposed method. It was, however, not rejected in either of the criteria in P3 for modified $k$-means and AAHC. Taken as a whole, the proposed method achieved a more precise estimation of time
TABLE 2 | Descriptive two one-sided tests (TOST) equivalence tests between ground truth TWs (time windows) and estimated TWs by the proposed consensus clustering (CC), modified k-means (MKMS), and atomize and agglomerate hierarchical clustering (AAHC) in individual subjects’ data from simulated ERP data.

| Comp-Meth | Cond | Criteria | p1  | p2  | DiffMu (ms) | EQ_interval (ms) |
|-----------|------|----------|-----|-----|-------------|-----------------|
| N2_CC     | C1   | Start    | 0.000 | 0.003 | 2.2         | 0.4             | 4.1             |
|           |      | End      | 0.000 | 0.003 | 2.6         | 1.0             | 4.1             |
|           |      | Duration | 0.000 | 0.000 | 0.4         | -1.6            | 2.3             |
| C2        | Start | 0.000 | 0.001 | 1.6 | -0.4 | 3.6 |
|           | End   | 0.000 | 0.002 | 2.3 | 0.7 | 4.0 |
|           | Duration | 0.000 | 0.000 | 0.7 | -1.2 | 2.6 |
| P3_CC     | C1   | Start    | 0.000 | 0.000 | 1.9         | 0.5             | 3.2             |
|           |      | End      | 0.018 | 0.000 | 2.8         | -4.7            | -0.9            |
|           |      | Duration | 0.380 | 0.000 | 4.7         | -6.8            | -2.6            |
| C2        | Start | 0.000 | 0.000 | 1.8 | 0.3 | 3.2 |
|           | End   | 0.008 | 0.000 | 2.6 | -4.4 | -0.7 |
|           | Duration | 0.266 | 0.000 | 4.3 | -6.4 | -2.2 |
| N2_MKMS   | C1   | Start    | 0.000 | 0.001 | 2.1         | 0.3             | 3.9             |
|           |      | End      | 0.000 | 0.041 | 3.5         | 1.9             | 5.1             |
|           |      | Duration | 0.000 | 0.000 | 1.4         | -0.3            | 3.1             |
| C2        | Start | 0.000 | 0.001 | 1.4 | -0.6 | 3.4 |
|           | End   | 0.000 | 0.030 | 3.5 | 2.0 | 5.0 |
|           | Duration | 0.000 | 0.001 | 2.1 | 0.5 | 3.7 |
| P3_MKMS   | C1   | Start    | 0.000 | 0.543 | 5.1         | 2.8             | 7.5             |
|           |      | End      | 0.548 | 0.000 | 5.1         | -7.3            | -3.0            |
|           |      | Duration | 0.997 | 0.000 | 10.3        | -13.8           | -6.7            |
| C2        | Start | 0.000 | 0.554 | 5.1 | 3.2 | 7.0 |
|           | End   | 0.654 | 0.000 | 5.5 | -7.8 | -3.1 |
|           | Duration | 0.999 | 0.000 | 10.6 | -14.0 | -7.3 |
| N2_AAHC   | C1   | Start    | 0.000 | 0.001 | 1.8         | 0.0             | 3.5             |
|           |      | End      | 0.000 | 0.104 | 4.1         | 2.7             | 5.5             |
|           |      | Duration | 0.000 | 0.000 | 2.3         | 0.9             | 3.7             |
| C2        | Start | 0.000 | 0.000 | 1.3 | -0.6 | 3.2 |
|           | End   | 0.000 | 0.039 | 3.6 | 2.1 | 5.1 |
|           | Duration | 0.000 | 0.001 | 2.3 | 0.8 | 3.9 |
| P3_AAHC   | C1   | Start    | 0.000 | 0.162 | 4.2         | 2.7             | 5.7             |
|           |      | End      | 0.244 | 0.000 | 4.3         | -6.2            | -2.4            |
|           |      | Duration | 0.999 | 0.000 | 8.5         | -10.6           | -6.4            |
| C2        | Start | 0.000 | 0.167 | 4.2 | 2.6 | 5.8 |
|           | End   | 0.276 | 0.000 | 4.4 | -6.3 | -2.6 |
|           | Duration | 0.999 | 0.000 | 8.6 | -10.7 | -6.6 |

Bold marked represent nonsignificant results. Comp-Meth, component of interest and the method; Cond, condition; C1, condition 1; C2, condition 2; p1, p value of lower bound; p2, p value of upper bound; DiffMu, difference of mean of two sets; and EQ_interval, confident equivalence interval.

Results of Real ERP Data

Time Windows and Topographies for ERPs of Interest

The clustering results (randomly selected) from running the proposed method on real data for N300 and prospective positivity components, the corresponding topography maps, and the spatial correlation of time-points are illustrated in Figure 8. N300 identified by the cluster maps 1 and 2 in the RS group, is illustrated by the colored area in Figure 8A for both PM and ongoing tasks. Furthermore, N300 identified by cluster map 1 in the HC group and two tasks (PM and ongoing), is illustrated in Figure 8B. Similarly, the prospective positivity component is isolated by the cluster maps 6 and 5 in the RS group for PM and ongoing tasks, respectively (Figure 8C). The identified prospective positivity by cluster maps 4 and 5 in the HC group for PM and ongoing tasks are illustrated, respectively (Figure 8D). The average topographies shown in Figure 8 are obtained from the selected time windows identified by the cluster maps. Hence, the statistical power analysis revealed that HC was
characterized by a more negative potential over the occipital-central electrodes \((p < 0.001)\). Additionally, a silently larger positive potential was localized over frontal-central electrodes compared to the RS group in N300. Moreover, a slightly more negative potential was observed over occipital-central electrodes \((p < 0.001)\) in the ongoing task from both RS and HC groups in the N300 component. Our results revealed no significant difference for prospective positivity regarding group factor; however, a larger positive potential was localized over central electrodes \((p < 0.0001)\) in the ongoing task comparing to the PM task.

**Statistical Analysis and Stability Test Results**

The mean \(p\) value and standard deviation (SD) were obtained from over 50 independent runs of the studied clustering methods and statistical analysis on the individual data (Table 3). Rendering to stability analysis, the proposed method (SD = 0.003) was more stable compared to modified \(k\)-means (SD = 0.006) for the main effect of group and less stable than AAHC (SD = 0.002) for N300 component. Interestingly, it was the most stable method compared to other studied clustering methods for both the main effect of task (SD = 0.002) and interaction between group and task (SD = 0.043). Besides, the statistical power analysis results showed that the main effects of group and task by the proposed method were significant \((p < 0.002\) for both factors). Likewise, the main effect of group was significant by the modified \(k\)-means \((p < 0.017)\) and AAHC \((p < 0.004)\). The main effect of task, however, was significant only via AAHC \((p < 0.013)\). Meanwhile, the interaction between group and task was not significant in
both modified k-means and AAHC. Similarly, the proposed method was statistically the most stable for the interaction between group and task (SD = 0.011) comparing to other studied clustering methods in prospective positivity (Table 3). Additionally, the main effect of task was significant (p < 0.0001), and, more importantly, the interaction between group and task was also significant (p < 0.007) by the proposed method. However, the main effect of group was not significant by the proposed method. The main effect of task was also significant by modified k-means (p < 0.0001) and AAHC (p < 0.0001), whereas, the main effect in group, and the interactions between group and task were not significant by both modified k-means and AAHC methods.

**DISCUSSION**

This study proposed a new methodology based on multisubject consensus clustering on spatiotemporal ERP data for the suitable time-window determination. To this end, we designed the stabilized multisubject consensus clustering in two levels described as follows: (i) subject resolution in which the stabilized consensus clustering was used to combine the results of various clusterings on each subject’s data in the group; (ii) group resolution in which the most suitable clustering for each group was obtained by consensus clustering of the clustering results of individual data. From the ERP technique point of view, the researchers using the ERP technique for the cognitive
neuroscience research often face up the challenge to determine a time window for an ERP, since the most popular textbook of ERP recommends the readers averaging the amplitudes in the time window as the measurement of the ERP peak amplitude (Luck, 2014). In terms of previous publications, we found that the determination of such a time window has mostly...
TABLE 3 | Mean p value and standard deviation (SD) calculations of statistical power analysis results in over 50 runs of study clustering methods on the individual data for the real data.

| Method          | N300 Group | Task | intGrTsk | Pros.Pos. Group | Task | intGrTsk |
|-----------------|------------|------|----------|-----------------|------|----------|
| Proposed (p value) | 0.002 0.002 0.058 | 0.590 0.000 0.007 | 0.004 0.013 0.145 | 0.662 0.000 0.246 |
| SD              | 0.003 0.002 0.043 | 0.227 0.000 0.011 | 0.006 0.075 0.225 | 0.199 0.000 0.156 |
| MKMS(p value)   | 0.017 0.101 0.303 | 0.614 0.000 0.150 | 0.004 0.013 0.145 | 0.662 0.000 0.246 |
| SD              | 0.004 0.013 0.145 | 0.662 0.000 0.246 | 0.002 0.009 0.133 | 0.201 0.000 0.131 |

IntGrTsk, interaction between group and task; Pros.Pos., prospective positivity. Bold marked represent significant results.

relied on the visual inspection, which can be subjective and bring bias to conclusions and difficulty for the readers to repeat the experiment. Therefore, the main objective of this work was to provide a reliable clustering-based mechanism (objective approach) for studying the temporal dynamic and sensory information about the subjects (i.e., brain responses). This was accomplished with the multilevel clustering mechanism and the time-window determination method. The clustering result from entire subjects entails important information about group response which is critical for studying the cognitive processes in ERP.

One issue in processing individual data is, apart from the need for sufficient trials for obtaining reliable ERP (Boudewyn et al., 2018) and the variety of brain responses in the trials, the variability associated with individual subjects’ brain responses, which is observable when ERPs are used to assess cognitive functions. The underlying assumption is that the variety in the trials and subject responses are involved in ERP, although in the ERP techniques, the assumption is that the ERP is phase-locked and time-locked. Therefore, each subject grants value to the statistical test in terms of differences between conditions or groups, which is through the variance across subjects assisting in the ability to detect a significant experimental effect (Kappenman and Luck, 2012a). Yet, in the literature, the individual responses were mostly addressed by fitting the cluster maps of individual data to the cluster maps of group average data (Murray et al., 2008; Koenig et al., 2014; Michel and Koenig, 2018; Berchio et al., 2019; Ruggeri et al., 2019). To cover this gap, we strove to cluster individual subject data in the first level and map the entire individual clusterings into a group as the ultimate clustering.

From the cluster analysis view of point, the various clustering strategies such as using the single clustering method on the different types of datasets; repeated clustering with a single clustering method and combining the results; and the multilevel-clustering methods applied to the individual dataset potentially affect the clustering quality (Abu-Jamous et al., 2013, 2015b; Liu et al., 2015; von Wegner et al., 2018). To investigate this issue and reliably feeding consensus clustering, two data-driven based mechanisms were appropriated before multilevel cluster analysis. First, consensus clustering configuration was performed aim to find the appropriate clustering methods. This was recognized by calculating the similarity between candidate clustering methods and modified k-means (benchmark) from individual data. Second, the stabilized clusterings were carried out by stabilizing the stochastic clusterings. Taken as a whole, these two procedures can make an additional sense of obtaining reliable and stable results instead of using a single clustering method or the conventional consensus clustering platform. Noteworthy to mention that clustering selection and stabilization can result in different configurations for various ERP data.

In accordance with the obtained results, two major differences were noticed between the proposed method and conventional clustering methods:

(i) The statistical test in this study revealed that the proposed method estimates a more precise time windows for individual subjects in comparison with the other conventional clustering methods in simulated data for both ERPs of interest (N2 and P3). The foremost reason is that our method uses the strength of multiple clustering methods and data-driven processing individual subject data to fit the suitable time windows for each condition/group, despite with using spatial consistency comparison between ERPs of individual and grand average data (Habermann et al., 2018; Michel and Koenig, 2018; Berchio et al., 2019).

(ii) According to the statistical analysis results (Table 3), the proposed method outperformed other benchmark methods regarding achieving more stability in the real data. Over 50 independent runs of the clustering on the same datasets, the estimation of the proposed method was with a much smaller variance. This indicates that the estimation of the ERP time window was much closer to the ground truth time window of the ERP, in contrast to the other methods. Such results from the real data also correspond to the ones from the simulation data, i.e., the estimation of the time window of an ERP was more accurate by the proposed method. Therefore, the results of the current study, based on analyzing the brain dynamics from the stimuli onset to the brain response, successfully explored the attention effect on the neural responses from the subjects in real ERP data.

The drawback of the proposed method, however, is that if the real ERP component is still embedded in ERP waveforms the determination of the time window of an ERP component cannot be precise. Indeed, this also happens in the visual inspection method to determine the time window of an ERP. Therefore, in order to determine the time window of an ERP component more precisely, the EEG preprocessing is very critical. The better the preprocessing is, the more precise and objective determination of the time window of an ERP is carried out in terms of the proposed method.

The results of analyzing the brain dynamic responses revealed that the brain electrical dynamics in obtained time windows were comparatively different in time-window properties (start, end, and duration) for different conditions/groups. Therefore,
from the clinical point of view, the brain responses from two
groups (RS and HC) to the stimuli onset were investigated
to identify N300 and prospective positivity components. This
can be interpreted as the fact of the variety of brain
response for the subjects in different condition/group. In
N300 component isolation, for example, the difference was
shown in cluster maps 1 (i.e., between RS and HC groups)
in PM tasks. Likewise, the duration differed in cluster maps
1 and 2 in the ongoing task between the groups. Again,
at the source level, a silently larger negative response was
observed in ongoing than PM task in both RS and HC
groups. These results demonstrate that RS patients with
schizophrenia showed a functional recovery of PM cue
detection during the event-based PM task. Consequently, the
electrophysiological data revealed the ability of symptomatically
remitted patients with schizophrenia to distinguish the PM
task from the ongoing task. This was reflected by the
significant main effect of task type among these two groups.
As a result, this finding showed a complementary viewpoint
to the prior studies (Fukumoto et al., 2014; Chen et al.,
2015). Our results can be employed for interpreting the
advantage of the treatment in RS patients in terms of
measuring/identifying the difference in ERPs of interest in
the observations. Therefore, this may indicate a degree of
functional recovery of preparatory attentional processes that
helps the processing of PM task in these subjects (RS patients)
during clinical remission. Thus, providing further evidence for
the recent researches demonstrating symptomatic remission in
schizophrenia is associated with a degree of functional recovery
of attentional processes.

CONCLUSION AND FUTURE WORKS

This work presents a multisubject consensus clustering
technique to explore spatiotemporal ERP by extracting
group-level information from individual responses. Our
proposed methodology has successfully extended the previous
research findings (Murray et al., 2008; Koenig et al., 2014;
Michel and Koenig, 2018) of cluster analysis of EEG/ERP.
Noteworthy to mention that we have proposed the multiset
consensus clustering method in the present study which
can work better for the group-level analysis. Since the
proposed method is not limited to just ERP data, it is very
interesting to apply the proposed method on other brain
imaging modalities for investigating the various types of
brain dynamics. Furthermore, the proposed method can also
be used as an appropriate tool to analyze the single-trial
EEG by considering suitable roles for the trials in higher
resolution (single-trials) in the future. Taken together, this
work emphasizes that, in the time-window determination
from spatiotemporal ERP, the temporal dynamics can be
extremely influenced through the measurement interval. It
is noteworthy that this methodology can be investigated on
different levels (i.e., groups, subjects, trials). The current
study also highlights that the obtained time windows are
sensitive to the responses from the subjects, which can
provide a better sense of understanding in information
processing of the neural responses. In order to show the
effectiveness of the proposed method, we have used the
simulated ERP dataset and the real ERP dataset. Indeed, the
selection of the real ERP dataset does not mean that the
proposed method only works for such an attention-related
ERP experiment. The proposed method has no limitation
on the experiment types of ERPs. Thereby, a toolbox
has been developed under the MATLAB platform, named
ERP_CC.2 Taken as a whole, we can rely on the information
retrieved by the new method, which reflects the attention
mechanism regarding the response to the stimuli in the
real data. We therefore believe that the EEG neuroimaging
method can be studied by the proposed methodology in
various dimensions to accomplish useful results in cognitive
neuroscience studies.

DATA AVAILABILITY STATEMENT

The datasets analyzed in this article are not publicly available.
Requests to access the datasets should be directed to GC.

ETHICS STATEMENT

The studies involving human participants were reviewed and
approved by the ethics committee of PLA general hospital. The
patients/participants provided their written informed consent to
participate in this study.

AUTHOR CONTRIBUTIONS

RM designed the methodology, implemented the algorithms,
performed data analysis, produced tables and figures, and wrote
the manuscript. YL performed statistical analysis, collected real
data, and analyzed the results. WD collected the real data. RF
prepared the simulated data. AN designed the methodology
and contributed to the final manuscript. GC designed the
study (real data), collected real data, and analyzed the results.
FY designed the study, analyzed the results, and wrote the
manuscript. All authors contributed to the article and approved
the submitted version.

FUNDING

This work was supported by National Natural Science
Foundation of China (Grant No. 91748105), National
Foundation in China (No. JCKY2019110B009) and
the Fundamental Research Funds for the Central
Universities (DUT2019) in Dalian University of Technology.

2https://doi.org/10.5281/zenodo.3541576
REFERENCES

Abu-Jamous, B., Fa, R., and Nandi, A. K. (2015a). Integrative Cluster Analysis in Bioinformatics. Copyright © 2015. Hoboken, NJ: John Wiley & Sons, Ltd, doi: 10.1002/978111906545

Abu-Jamous, B., Fa, R., Roberts, D. J., and Nandi, A. K. (2015b). UNCLES: method for the identification of genes differentially consistently co-expressed in a specific subset of datasets. BMC Bioinformatics 16:184. doi: 10.1186/s12859-015-0614-0

Abu-Jamous, B., Fa, R., Roberts, D. J., and Nandi, A. K. (2013). Paradigm of tunable clustering using binarization of consensus partition matrices (Bi-CoPaM) for gene discovery. PLoS One 8:e0056432. doi: 10.1371/journal.pone.0056432

Bailey, N. W., Freedman, G., Raj, K., Sullivan, C. M., Rogasch, N. C., Chung, S. W., et al. (2019). Mindfulness meditators show altered distributions of early and late neural activity markers of attention in a response inhibition task. PLoS One 14:e0203096. doi: 10.1371/journal.pone.0203096

Bérchio, C., Küng, A.-L., Kumar, S., Cordera, P., Dayer, A. G., Aubry, J.-M., et al. (2019). Eye-gaze processing in the broader bipolar prototype revealed by electrical neuroimaging. Psychiatry Res. 291, 42–51. doi: 10.1016/j.psychres.2019.07.007

Bezdek, J. C. (1981). Pattern Recognition With Fuzzy Objective Function Algorithms. Berlin: Springer, doi: 10.1007/978-1-4757-0450-1

Boudewyn, M. A., Luck, S. J., Farrens, J. L., and Kappenman, E. S. (2018). How many trials does it take to get a significant ERPs effect? It depends. Psychophysiology 55:e13049. doi: 10.1111/psyp.13049

Cacioppo, S., Weiss, R. M., Runesha, H. B., and Cacioppo, J. T. (2014). Dynamic thought, “in Machinery of the Mind: Data, Theory, and Speculations About Brain Dynamics” in Higher Brain Function, eds E. R. John, T. Harmony, L. S. Prichep, M. Valdés-Sosa, and P. A. Valdés-Sosa (Boston, MA: Birkhäuser Boston), 209–224. doi: 10.1007/978-1-4757-1083-0_10

Chen, G., Zhang, L., Ding, W., Zhou, R., Xu, P., Lu, S., et al. (2015). Event-related brain potential correlates of prospective memory in symptomatically remitted male patients with schizophrenia. Front. Behav. Neurosci. 74, 239–248. doi: 10.3109/00207459408987242

Kappenman, E. S., and Luck, S. J. (2012b). “ERP components and selective attention,” in The Oxford Handbook of Event-Related Potential Components, eds S. J. Luck and E. S. Kappenman (Oxford: Oxford University Press), 3–30. doi: 10.1093/oxfordhb/9780195374148.013.0014

ACKNOWLEDGMENTS

This study is to memorize TR for his great help to all the authors, especially FC, AN, and RM.
topographic analysis. Neuroimage 122, 65–72. doi: 10.1016/j.neuroimage.2015.07.080

Mahini, R., Xu, P., Chen, G., Li, Y., Ding, W., Zhang, L., et al. (2019). Optimal number of clusters by measuring similarity among topographies for spatio-temporal ERP analysis. arXiv [Preprint]. Available online at: https://zenodo.org/badge/latestdoi/197869407

Mahini, R., Zhou, T., Li, P., Nandi, A. K., Li, H., Li, H., et al. (2017). “Cluster Aggregation for generating event-related potentials,” in Advances in Neural Networks - ISNN 2017. ISNN 2017. Lecture Notes in Computer Science, eds F. Cong, A. Leung, and Q. Wei (Cham: Springer International Publishing), 507–515. doi: 10.1007/978-3-319-59081-3_59

Meila, M. (2007). Comparing clusterings - an information based distance. J. Mult. Anal. 98, 873–895. doi: 10.1016/j.jmva.2006.11.013

Micah, M. M., Lucia, M. D., Brunet, D., and Michel, C. M. (2009). Principles of Topographic Analyses for Electrical Neuroimaging. Cambridge, MA: MIT Press, doi: 10.7551/mitpress/9780262013086.003.0002

Michel, C. M., and Koenig, T. (2018). EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. NeuroImage 180, 577–593. doi: 10.1016/j.neuroimage.2017.11.062

Monti, S., Tamayo, P., Mesirov, J., and Golub, T. (2003). Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Mach. Learn. 52, 91–118. doi: 10.1023/a:1023949509487

Mu, Y., and Han, S. (2010). Neural oscillations in self-referential processing. NeuroImage 53, 757–768. doi: 10.1016/j.neuroimage.2010.07.008

Murray, M. M., Brunet, D., and Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. Brain Topogr. 20, 249–264. doi: 10.1007/s10548-008-0054-5

Nguyen, N., and Caruana, R. (2007). “Consensus clusterings,” in Seventh IEEE International Conference on Data Mining (ICDM 2007). Omaha, NE, 607–612. doi: 10.1109/ICDM.2007.73

Pascal-Marqui, R. D., Michel, C. M., and Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans. Biomed. Eng. 42, 658–665. doi: 10.1109/30.491164

Pena, J. M., Lozano, J. A., and Larranaga, P. (2019). An empirical comparison of four initialization methods for the k-means algorithm. Pattern Recogn. Lett. 20, 1027–1040. doi: 10.1016/S0167-8655(99)00069-0

Pourtois, G., Delplaque, S., Michel, C., and Vuilleumier, P. (2008). Beyond conventional event-related brain potential (ERP): exploring the time-course of visual emotion processing using topographic and principal component analyses. Brain Topogr. 20, 265–277. doi: 10.1007/s10548-008-0053-6

Rogers, J. L., Howard, K. I., and Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental groups. Psychol. Bull. 113, 553–565. doi: 10.1037/0033-2909.113.3.553

Rotshtein, P., Richardson, M. P., Winston, J. S., Kiebel, S. J., Vuilleumier, P., Eimer, M., et al. (2010). Amygdala damage affects event-related potentials for fearful faces at specific time windows. Hum. Brain Mapp. 31, 1089–1105. doi: 10.1002/hbm.20921

Ruggeri, P., Meziane, H. B., Koenig, T., and Brandner, C. (2019). A fine-grained time course investigation of brain dynamics during conflict monitoring. Sci. Rep. 9:3667. doi: 10.1038/s41598-019-04277-3

Sawaki, R., Geng, J. J., and Luck, S. J. (2012). A common neural mechanism for preventing and terminating the allocation of attention. J. Neurosci. 32, 10725–10736. doi: 10.1523/JNEUROSCI.1846-12.2012

Seaman, M. A., and Serlin, R. C. (1998). Equivalence confidence intervals for two-group comparisons of means. Psychol. Methods 3, 403–411. doi: 10.1037/1082-989X.3.4.403

Sipola, T., Cong, F., Ristaniemi, T., Alluri, V., Toivainen, P., Brattico, E., et al. (2013). “Diffusion map for clustering fMRI spatial maps extracted by independent component analysis,” in 2013 IEEE International Workshop on Machine Learning for Signal Processing (MLSP), Southampton, 1–6. doi: 10.1109/MLSP.2013.6661923

Song, Y., Zhang, Z., Hu, T., Gong, X., and Nandi, A. K. (2019). “Identify of spatial similarity of electroencephalography (EEG) during working-memory maintenance,” in 2019 27th European Signal Processing Conference (EUSIPCO). A Coruña, 1–5. doi: 10.23919/EUSIPCO.2019.8902595

Strehl, A., and Ghosh, J. (2003). Cluster ensembles - a knowledge reuse framework for combining multiple partitions. J. Mach. Learn. Res. 3, 583–617. doi: 10.1162/153244303231877335

Tibshirani, R., and Walther, G. (2005). Cluster validation by prediction strength. J. Comput. Graph. Stat. 14, 511–528. doi: 10.1198/106186005X9243

Tzovara, A., Murray, M. M., Plomp, G., Herzog, M. H., Michel, C. M., and De Lucia, M. (2012). Decoding stimulus-related information from single-trial EEG responses based on voltage topographies. Pattern Recogn. 45, 2109–2122. doi: 10.1016/j.patcog.2011.04.007

Van Overwalle, F., Van den Eede, S., Baetens, K., and Vandekerckhove, M. (2009). Trait inferences in goal-directed behavior: ERP timing and localization under spontaneous and intentional processing. Soc. Cogn. Affect. Neurosci. 4, 177–190. doi: 10.1093/scan/nsp003

von Wegner, F., Knaut, P., and Laufs, H. (2018). EEG microstate sequences from different clustering algorithms are information-theoretically invariant. Front. Comput. Neurosci. 12:70. doi: 10.3389/fncom.2018.00070

Williams, N., Nasuto, S. J., and Saddy, J. D. (2015). Method for exploratory cluster analysis and visualisation of single-trial ERP ensembles. J. Neurosci. Method 250, 22–33. doi: 10.1016/j.jneumeth.2015.02.007

Wills, A. J., Lavric, A., Hemmings, Y., and Surrey, E. (2014). Attention, predictive learning, and the inverse base-rate effect: evidence from event-related potentials. Neuroimage 87, 61–71. doi: 10.1016/j.neuroimage.2013.05.060

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Mahini, Li, Ding, Fu, Ristaniemi, Nandi, Chen and Cong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.