

# Combining self-organizing mapping and supervised affinity propagation clustering approach to investigate functional brain networks involved in motor imagery and execution with fMRI measurements

## OPEN ACCESS

### Edited by:

Mingzhou Ding,  
University of Florida, USA

### Reviewed by:

Xiaotong Wen,  
Renmin University of China, China  
Yingchun Zhang,  
University of Houston, USA

### \*Correspondence:

Jiang Zhang and Qi Liu,  
Department of Medical Information  
Engineering, School of Electrical  
Engineering and Information, Sichuan  
University, No. 24, South Section 1,  
Yihuan Road, Chengdu 610065, China  
zhangjiang\_@hotmail.com;  
liuqi@scu.edu.cn  
HuaFu Chen,  
School of Life Science and  
Technology, University of Electronic  
Science and Technology of China,  
No. 4, Section 2, North Jianshe Road,  
Chengdu 610054, China  
chenhf@uestc.edu.cn

**Received:** 06 March 2015

**Accepted:** 29 June 2015

**Published:** 17 July 2015

### Citation:

Zhang J, Liu Q, Chen H, Yuan Z,  
Huang J, Deng L, Lu F, Zhang J,  
Wang Y, Wang M and Chen L (2015)  
Combining self-organizing mapping  
and supervised affinity propagation  
clustering approach to investigate  
functional brain networks involved in  
motor imagery and execution with  
fMRI measurements.  
*Front. Hum. Neurosci.* 9:400.  
doi: 10.3389/fnhum.2015.00400

**Jiang Zhang**<sup>1\*</sup>, **Qi Liu**<sup>1\*</sup>, **HuaFu Chen**<sup>2\*</sup>, **Zhen Yuan**<sup>3</sup>, **Jin Huang**<sup>4</sup>, **Lihua Deng**<sup>1</sup>,  
**Fengmei Lu**<sup>3</sup>, **Junpeng Zhang**<sup>1</sup>, **Yuqing Wang**<sup>5</sup>, **Mingwen Wang**<sup>6</sup> and **Liangyin Chen**<sup>7</sup>

<sup>1</sup> Department of Medical Information Engineering, School of Electrical Engineering and Information, Sichuan University, Chengdu, China, <sup>2</sup> School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China, <sup>3</sup> Bioimaging Core, Faculty of Health Sciences, University of Macau, Macau, China, <sup>4</sup> School of Foreign Studies, University of Electronic Science and Technology of China, Chengdu, China, <sup>5</sup> CAS Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety National Center for Nanoscience and Technology of China, Beijing, China, <sup>6</sup> School of Mathematics, Southwest Jiaotong University, Chengdu, China, <sup>7</sup> School of Computer Science, Sichuan University, Chengdu, China

Clustering analysis methods have been widely applied to identifying the functional brain networks of a multitask paradigm. However, the previously used clustering analysis techniques are computationally expensive and thus impractical for clinical applications. In this study a novel method, called SOM-SAPC that combines self-organizing mapping (SOM) and supervised affinity propagation clustering (SAPC), is proposed and implemented to identify the motor execution (ME) and motor imagery (MI) networks. In SOM-SAPC, SOM was first performed to process fMRI data and SAPC is further utilized for clustering the patterns of functional networks. As a result, SOM-SAPC is able to significantly reduce the computational cost for brain network analysis. Simulation and clinical tests involving ME and MI were conducted based on SOM-SAPC, and the analysis results indicated that functional brain networks were clearly identified with different response patterns and reduced computational cost. In particular, three activation clusters were clearly revealed, which include parts of the visual, ME and MI functional networks. These findings validated that SOM-SAPC is an effective and robust method to analyze the fMRI data with multitasks.

**Keywords:** motor imagery, motor execution, functional magnetic resonance imaging, self-organizing mapping, affinity propagation clustering

## Introduction

Motor imagery (MI) and motor execution (ME), have shown their potentials for the investigations of rehabilitation in movement disorders and brain-computer interfaces (BCIs) as well as for training athletes and musicians (Jeannerod, 1995; Porro et al., 1996; Pfurtscheller and Neuper, 2001; Malouin et al., 2004; Brouziyne and Molinaro, 2005; Kimberley et al., 2006; Lotze and Halsband, 2006; Felton et al., 2007; Chen et al., 2009). Now advances in brain imaging are undergoing a transition from mapping sites of cortical activations toward identifying the brain networks that connect these sites together into dynamic systems in time-frequency domain (Buxton et al., 2004; Ding et al., 2006; Yuan and Ye, 2013). As such, quantifications of brain networks have become a major field of interest in neuroimaging and neurointegration for the investigation of MI and ME (Ding et al., 2006; Chen et al., 2009; Wang et al., 2010; Yuan and Ye, 2013). For assessing functional connectivity in human brains involved in MI and ME, it is essential to conduct pattern clustering of complex brain networks using fMRI data.

Interestingly clustering analysis methods like self-organizing mapping (SOM) and affinity propagation clustering (APC) are able to reveal inter-subject differences in the temporal dynamics of the fMRI signals without a prior model (Goutte et al., 1999; Frey and Dueck, 2007; Liao et al., 2008; Ren et al., 2014). SOM (Kohonen, 1982, 1990, 1995; Haykin, 1999) transforms the incoming signal patterns into low dimensional discrete maps (Kohonen, 1990, 1995; Haykin, 1999) to generate time courses of a predetermined number of exemplars (Peltier et al., 2003). Meanwhile, APC utilizes a pattern cluster tool that is able to sort out exemplars from the whole data points, in which clusters of data points surrounding the specific exemplars are generated (Frey and Dueck, 2007). However, although APC is considered an efficient analysis method (Frey and Dueck, 2007; Mézard, 2007), it is very hard to use this scheme to conduct pattern clustering of complex functional networks using fMRI measurements because substantial storage and computer memory are required for clustering analysis if implemented with a personal computer (PC).

To improve the computational efficiency and resolve memory limitation problem, principal component analysis (PCA) was first used to narrow down the whole brain voxels before APC clustering analyses were performed for the remaining voxels (Zhang et al., 2011). And we name this updated scheme as PCA-APC. The fMRI datasets can be processed by PC using PCA-APC when the identified voxels by PCA are involved in the computation. However, when the number of data points is significantly increased, the memory insufficiency problem will appear again and in the end the PCA-APC scheme would fail to process the large datasets from the imaging volumes. In addition, APC is not able to achieve an automatic acquisition of the optimal cluster numbers, which will affect the accuracy and reliability of the established brain networks. In order to resolve these issues, a novel data-driven clustering method, namely SOM-SAPC that combines SOM and supervised affinity propagation clustering (SAPC), was proposed to directly perform pattern clustering of complex functional networks involved in MI and ME. And the

developed method has the capability to overcome the memory limitation problem when implemented on a PC.

## Materials and Methods

### Developed SOM-SAPC for Brain Network Analysis

To reduce the computation cost involved in the brain network analysis, SOM was first implemented for the analysis of fMRI data. Then the generated experimental data were further processed using SAPC to identify brain activation patterns. In this study, SOM adopts a correlation distance metric between the input vector and weight vectors to spot the winning neuron (Chuang et al., 1999; Ngan and Hu, 1999; Ngan et al., 2002; Peltier et al., 2003; Hausfeld et al., 2014), which is different from previous strategy that utilizes the Euclidean distance (Kohonen, 1982, 1990, 1995; Chuang et al., 1999; Haykin, 1999; Ngan and Hu, 1999; Ngan et al., 2002; Peltier et al., 2003; Hausfeld et al., 2014). When SOM is implemented, each generated exemplar will have one corresponding time course. Then further operations can be performed to find the correlation coefficient (CC) between the time course of each voxel of brain slices and the time courses of all exemplars. The biggest CC will occur when the voxel and its corresponding exemplar are of the same cluster.

To ensure the exemplar's time courses can sufficiently retain brain image information (Peltier et al., 2003; Liao et al., 2008), in this study the number of exemplars, 100 was adopted for Kohonen's SOM analysis of fMRI data. If one exemplar represents one cluster, 100 clusters would be involved for clustering analysis, which might bring difficulties in interpreting fMRI data. Therefore, improved SOM-based fMRI data analyses have resorted to further clustering based on the 100 clusters acquired so as to ascertain patterns of brain functions (Chuang et al., 1999; Ngan and Hu, 1999; Ngan et al., 2002; Peltier et al., 2003; Hausfeld et al., 2014).

Following SOM estimation, SAPC is further conducted for the intensity-normalized time courses of the 100 analysis exemplars. Additionally, because only these exemplars' time courses are subjected to SAPC analysis, the new scheme is able to significantly decrease the number of calculations that may otherwise be involved in clustering analysis. "Further clustering" (Peltier et al., 2003) usually necessitates a prior designated number of clusters. It is noted APC is able to generate the optimal set of exemplars and associated clusters (Zhang et al., 2011). However, only SAPC features an automatic acquisition of cluster numbers via supervising the index of clusters and optimizing the input preferences whereas APC is incapable of such automation.

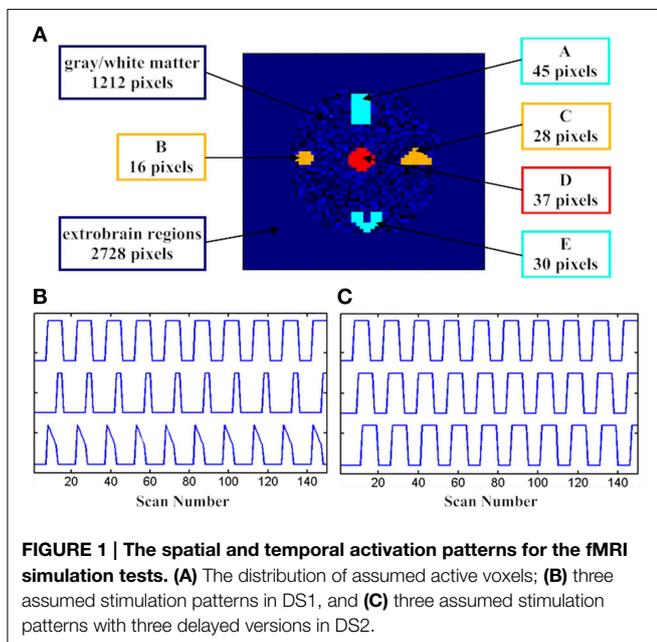
When the general APC is utilized to analyze the data points, the similarity between a pair of normalized vectors  $y_i$  and  $y_j$  is defined by  $s(i, j) = -\|y_i - y_j\|^2, i \neq j$ . Meanwhile, self-similarity  $s(i, i)$  is considered as an input preference ( $p$ ) (Frey and Dueck, 2007; Guan et al., 2011). In APC, the value of  $p$  as a constant can influence the clustering quality and the number of clusters. However, it is very difficult to specify beforehand the exact  $p$  because that case is most suitable for a given problem.

To search for an optimized  $p$ , the improved method SAPC has to be employed, which is able to directly supervise the Silhouette index of clusters (Kaufman and Rousseeuw, 1990; Zhang et al., 2011). The calculation of silhouette values is accomplished via the SILHOUETTE function of MATLAB (silhouette is plotted using the Euclidean distance). In particular, the averaged silhouette values represent the clustering quality, in which a larger value will generally provide a more accurate estimate of the clustering quality.

The Golden-section search (Antonioni and Lu, 2007) of the input preferences  $p$  is implemented in this study to generate an optimal set of clusters. In the Golden-section search, iterations are run until the desired accuracy for the maximum value of the mean silhouette is achieved (Antonioni and Lu, 2007). As a result, a better clustering quality is realized and the cluster number is also identified. Finally, the voxels for each slice of the brain corresponding to the 100 exemplars are clustered based on the clustering results of the 100 exemplars. The index of brain activity is defined as  $\lambda_i = \max(l(i)) - l(i)$  (Zhang et al., 2011), in which  $l(i)$  determines the Euclidean distance of the  $i$ th point with its associated cluster center, whereas the maximum  $\max(l(i))$  is subject to the same cluster.

### Generation of Simulated Data

Simulation tests based on the datasets generated with different hemodynamic responses were first conducted to examine the performance of the developed SOM-SAPC. For the test geometry, a composite image with 4096 voxels was generated on an axial brain. The simulated fMRI image consists of 5 active areas with 156 voxels. In addition, a time invariant margin that has 2728 voxels is added in the test geometry whereas a stochastic variant texture is specified for gray/white matter and ventricles that have 1212 voxels. The 5 brain areas won't have identical sizes and shapes as shown in **Figure 1A**.



The supposed box-car-like signals, DS1, DS2, and DS3, which are present in the 5 brain activity regions, are utilized here to define the brain activation patterns. DS1 is used to describe the change of the hemodynamic responses to different stimuli patterns. In addition, as displayed in **Figure 1B**, subregions “A\_E” (“A” and “E”), “B\_C” (“B” and “C”), and “D” have different temporal signals. The spatial pattern for “A\_E” was remote while its temporal pattern was accordant. Both the spatial and temporal patterns were combined together to generate a homogeneous active source for “A\_E” and we did the same thing for “B\_C” and “D”. Gaussian noise was also added to the time course of each voxel from the subregions with signal-to-noise ratio (SNR) equal to 1.0.

DS2 is able to show the timing variability between a stimulation paradigm and an activation delay of the hemodynamic response. The spatial pattern and SNR from DS2 were identical with those from DS1. Three temporal patterns with three delayed formula (delays of 0, 4, or 8 s; **Figure 1C**) of the “expected” box-car-like timing function are displayed in the first row of **Figure 1C**.

DS3 consisted of a given signal mixed with Gaussian noise, and the given signal was depicted in the first row of **Figure 1B**. DS3 was used to demonstrate the SNR variability. Five sources (subregions “A” and “E,” subregions “B” and “C” and subregion “D”) were correlated with the “expected” boxcar-like timing function. The differences in SNRs (1.2, 1.0, 0.8, 0.6, and 0.4) were computed using the integrated clustering approach.

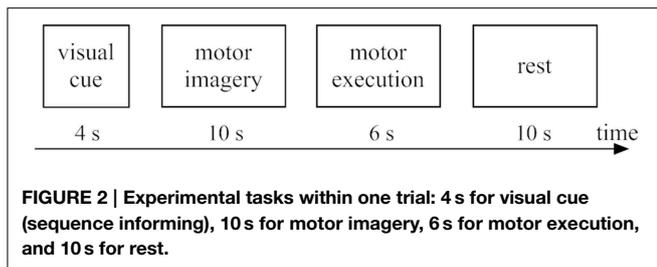
### Experimental Paradigm

#### Subjects

Ten right-handed subjects (aged: 19–25 years) were recruited to participate in the fMRI study. All the participants have normal visual acuity. Nobody has reported neurological diseases or mental disorders in the past 10 years. The subjects were assessed by the Edinburgh Handedness Inventory to exclude any possible left-handers. The clinical tests were consented by the subjects and the local Institutional Review Board of the West China Hospital of Sichuan University.

#### Tasks

The experimental design for the present work was the same as our previous fMRI studies (Chen et al., 2009). fMRI tests were performed with 2 conditions including bimanual MI and ME, which covered 10 trials (runs). The stimulus duration for each run was 30 s, as displayed in **Figure 2**. The stimuli with block design tasks started with sequence informing (4 s), and then followed by MI (10 s), ME (6 s), and resting (10 s). During each run, participants need be familiar with four sequentially pictures. And each sequentially picture is composed of a specific order of finger tapping task. Then, subjects began to perform finger tapping imagery tasks with the same order that was previously informed by the visual stimulus tasks. It should be noted the screen was kept black during 10 s period for MI. Then a new cue was shown on the screen for 6 s, which requested the subject to implement a finger tapping task (ME) with the same content as they had imaged. For our fMRI experimental design, the reason why MI was followed by ME is to make sure the subjects are



able to concentrate on the MI task. Before fMRI tests, we would train the participants so that they could perform the experiments very well.

### Data Acquisition

Experimental tests were conducted using a 3.0-T GE Signa scanner (MR Research Center of the West China Hospital of Sichuan University, Chengdu, China). The gradient-recalled echo planar imaging (EPI) sequence settings were as follows: 30 transverse slices, TR = 2000 ms, TE = 30 ms, FOV = 24 cm, acquisition matrix =  $64 \times 64$ , slice thickness = 5 mm (without gap), voxel size =  $3.75 \times 3.75 \times 5 \text{ mm}^3$ , and flip angle =  $90^\circ$ . We collected 155 images for the fMRI tests and the data acquisition time was about 310 s.

### Data Processing Procedure

Experimental data were processed using SPM8 software (<http://www.fil.ion.ucl.ac.uk/spm/>). In consideration of magnetization equilibrium and for the subjects to be familiar with the procedures (Chen et al., 2009), the first 5 images acquired would be rejected for each run while the additional 150 ones would be kept and calibrated using the methods developed in SPM8. In addition, realignment was adopted to eliminate the head movement. Spatial normalization was also implemented for the images based on the MNI EPI template. Further, fMRI data were smoothed with a FWHM of 8 mm. Ultimately signals captured in the brain areas were processed to validate SOM-SAPC.

## Results

### Simulation Tests

#### Compared with PCA

To compare with SOM, PCA is first adopted to generate the primary image before the SAPC analysis can be further implemented for the primary image. In PCA, the selected time courses and their numbers are determined by principal components and the selected threshold value (Zhang et al., 2011). In this study, to ensure the chosen principal components retain enough primitive information, the cumulative sum of the variances is set as 90%. For PCA (Zhang et al., 2011), out-of-memory won't occur when performed with a PC (Intel(R) Core (TM) i5 CPU @ 2.67 GHz, RAM 4 GB) for this simulation case. We found that compared with PCA, SOM enables a more efficient clustering analysis and the computational cost based on SOM method was also much lower, as displayed in **Table 1**. The disparity between the results after SOM and PCA analysis

was due to the fact that the quantity of data left over for SAPC analysis after SOM-based processing is far less than that after the PCA-based acquisition of the primary image (**Table 1**). The identified brain activity regions and the maximum memory utilized by SOM-SAPC were compared to that from PCA-SAPC when they were implemented to process datasets DS1 and DS2 (see **Figures 3A1–C**) for the reconstructed images and **Table 1** for the used maximum PC memory (MB). It was observed from **Figure 3** that all the five brain activity regions can be clearly identified by both SOM-SAPC and PCA-SAPC, and the Jaccard coefficient (*JC*) values of both SOM-SAPC and PCA-SAPC are equal to 1. However, it was also found from **Table 1** that less memory was required for the SOM-SAPC. In addition, the *JC* values and maximum memory required for different SNRs were also computed and compared between SOM-SAPC and PCA-SAPC for the analysis of datasets DS3 (see **Figures 3D,E**). We can see that under SNRs  $\geq 0.8$ , the *JC*-values of SOM-SAPC is almost equal to that from PCA-SAPC. However, when SNR = 0.6, the *JC*-value of SOM-SAPC is smaller when compared to that of PCA-SAPC. **Figure 3E** also showed that SOM-SAPC has smaller memory requirements for all the cases with various SNRs.

### Compared with other Clustering Analyses

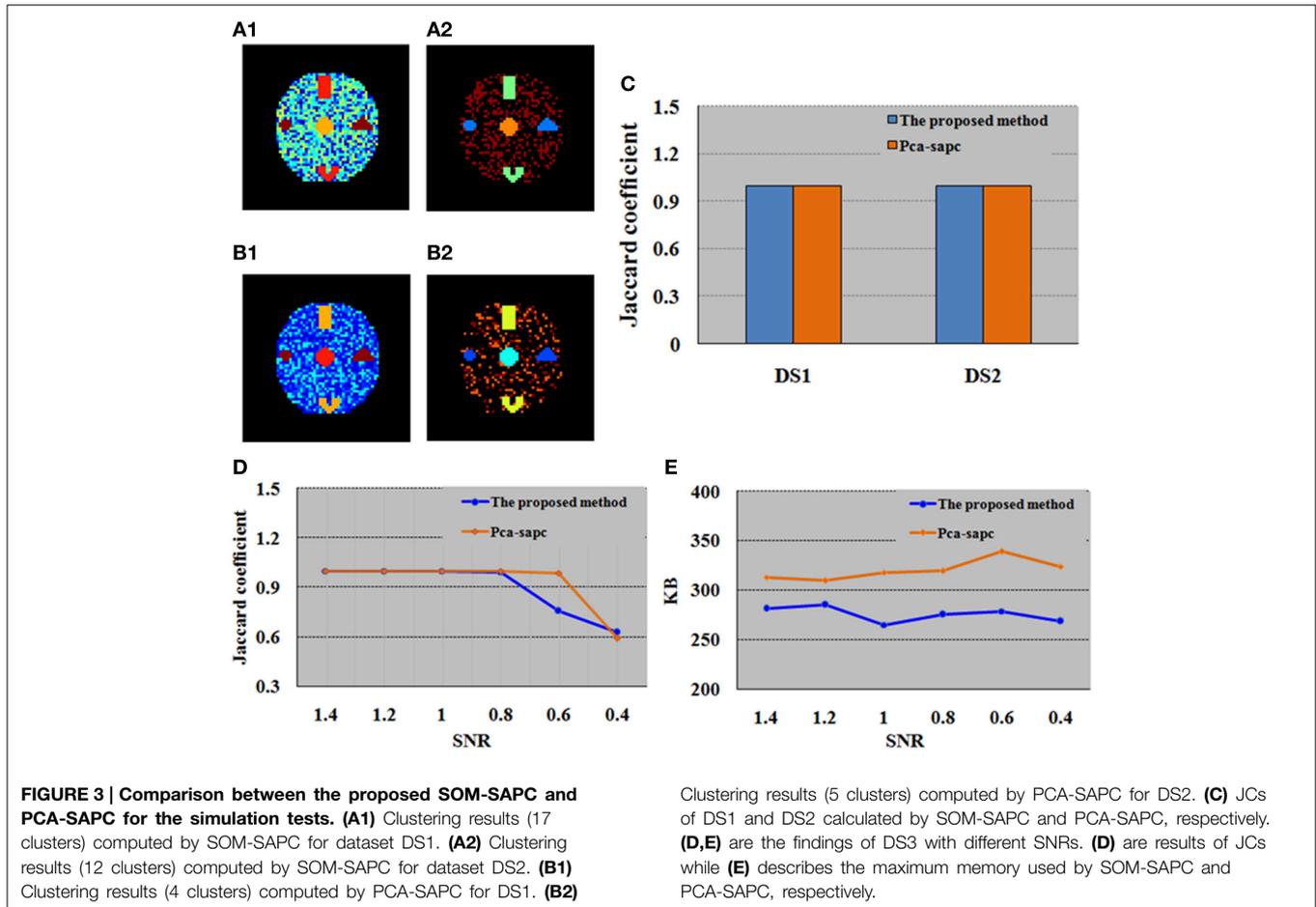
Simulated data were analyzed using different clustering methods to validate and show the advantages of the developed SOM-SAPC. The image analysis results are provided in **Figures 4–7**, in which each color represents a unique cluster for (c1), (c2), (d1), (d2), and (e), respectively. The *JC* is employed to examine the performance of the cluster (Anderberg, 1973; Baumgartner et al., 1998; Dimitriadou et al., 2004) and *JC* is equal to  $a/(a+b+c)$ . It is noted here *a*, *b*, and *c* represents the number of true-positive findings (TP), the number of false-negative findings (FN), and number of false-positive findings (FP), respectively.

**Figures 4A, 5A** show the exemplar matrix (100 nodes) analysis results calculated from SOM. We found that SOM is capable of identifying the winner nodes. In particular, the exemplar time course was also generated for DS1 and DS2 based on SOM.

**Figures 4C1–E, 5C1–E** display the clustering results produced from SOM-SAPC, *k*-centers clustering (MacQueen, 1967) and hierarchical clustering (HC) analysis in DS1 and DS2, respectively (the distance measured between two data points is the Euclidean distance, and the hierarchical cluster tree is calculated using the shortest distance) (Goutte et al., 1999; Chen et al., 2006; Liao et al., 2008; Ren et al., 2014). As plotted in **Figure 4**, the hemodynamic responses to different stimulus patterns are clearly identified for DS1. For example, **Figure 4C1** displays the analysis results calculated by SOM-SAPC, which generates 17 clusters. Activated subregions “A\_E,” “B\_C” and “D” were also reorganized to generate three specific clusters. **Figure 4C2** plots the imaging results reconstructed when the number of neighborhood voxels within a cluster is over five with SOM-SAPC. With the same number of clusters as in SOM-SAPC, the results from the *k*-centers clustering analysis with the max and min of the 50 *JC*-values [for the 50 repetitions of the method, because *k*-centers clustering often needs to rerun many times with different initializations in an attempt to find a good

**TABLE 1 | Discrepancies between PCA and SOM-based analysis via Matlab.**

| Dataset | Method | Number of time courses left over for SACP | Computation time of SACP (seconds) | Maximum memory utilized (MB) |
|---------|--------|---|------------------------------------|------------------------------|
| DS1     | SOM    | 100                                       | 30.7206                            | SOM-SACP                     |
|         | PCA    | 519                                       | 146.0798                           | PCA-SACP                     |
| DS2     | SOM    | 100                                       | 30.9426                            | SOM-SACP                     |
|         | PCA    | 432                                       | 108.2746                           | PCA-SACP                     |

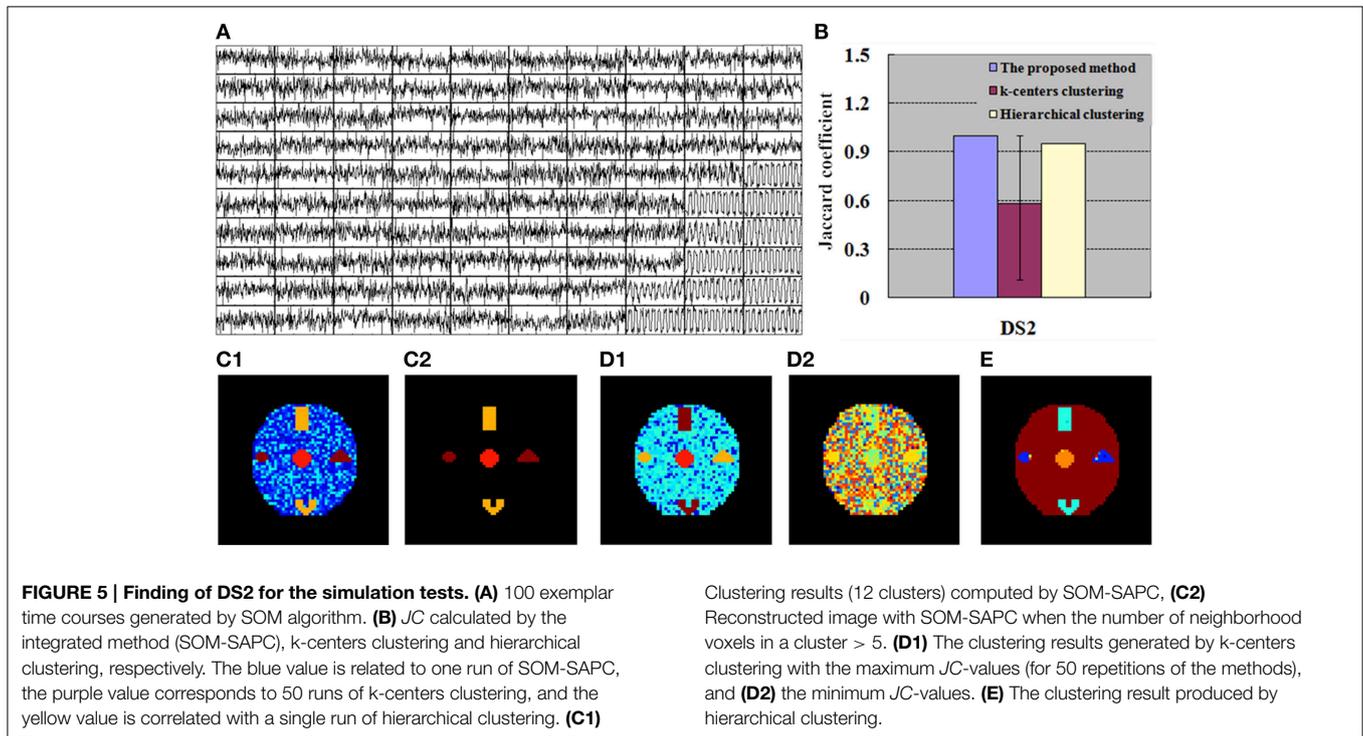
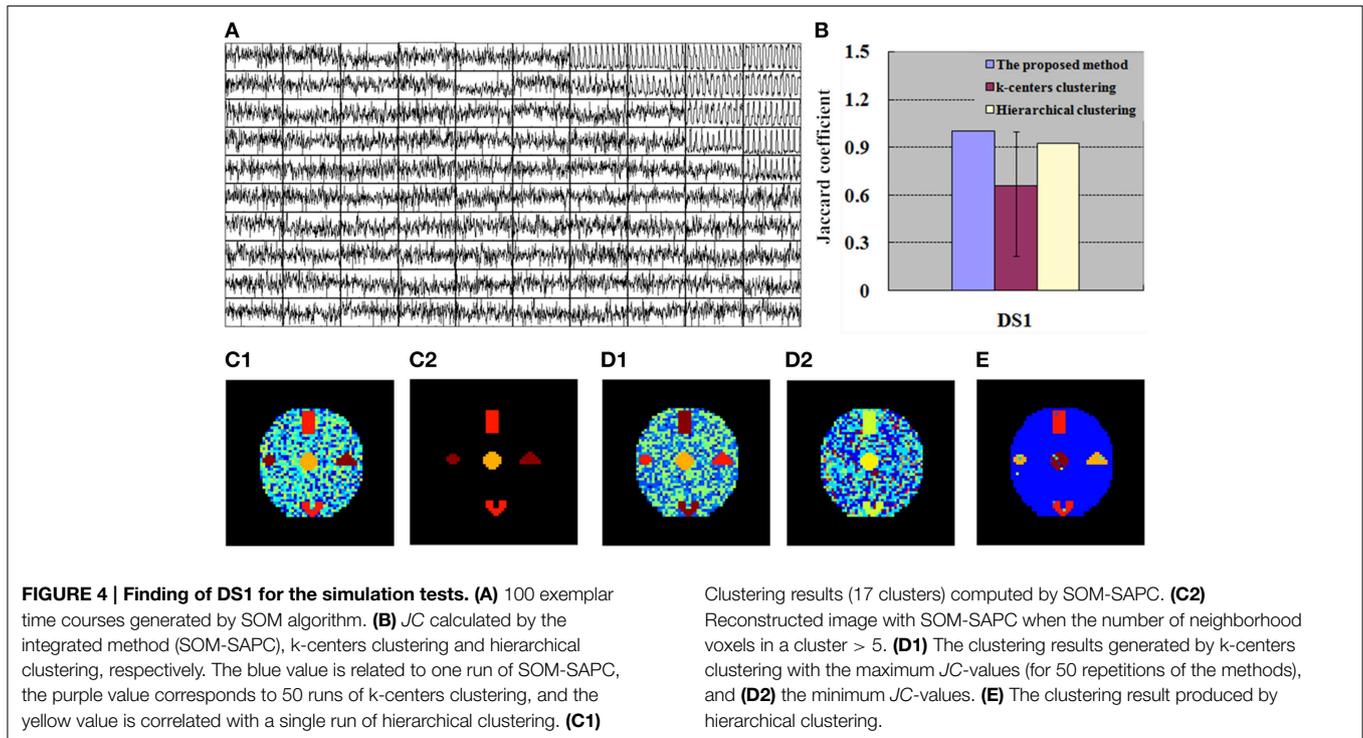


solution (Frey and Dueck, 2007)] are plotted in **Figures 4D1,D2**, respectively while the clustering result of the HC is given in **Figure 4E**. Likewise, the results of DS2, illustrating timing variability between the stimulation paradigm and hemodynamic response (activation delay) with the same number of clusters as that of the proposed method, are provided in **Figure 5**.

We then implemented *JC* analysis to provide a better comparison for the clustering capabilities among the different methods (shown in **Figures 4B, 5B**). It is noted from **Figures 4B, 5B** that the blue column represents the *JC* in the proposed method, the purple column is the *JC* of k-centers clustering, and the yellow column is the *JC* calculated by the HC method. We found from **Figures 4, 5** that SOM-SACP is able to clearly identify the number of clusters. With the same number of

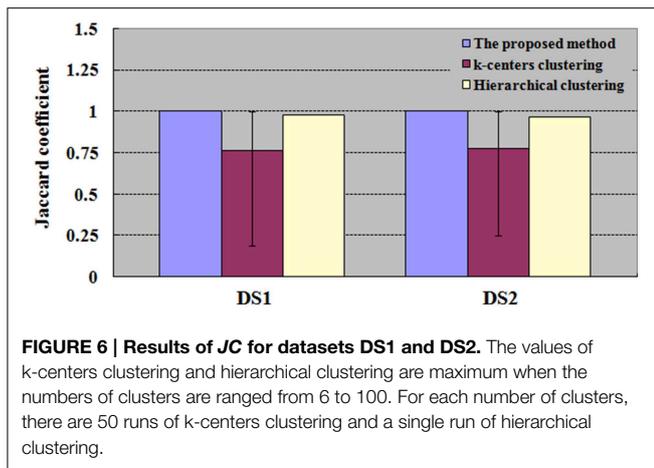
clusters, the computed results using k-centers clustering and HC were also obtained. In addition **Figures 4B, 5B** show the *JC* performance parameters. We observed from **Figures 4, 5** that in both SOM-SACP and HC analyses, each column *JC*-value is generated by a single run, while in k-center clustering, each column *JC*-value is generated by 50 runs. The large *JC*-values are obtained by SOM-SACP, and pixels in the activated subregions “A\_E,” “B\_C” and “D” are correctly grouped into three different clusters.

With the same number of clusters as in the integrated method (SOM-SACP), **Figures 4B, 5B** plot the *JC* performance. Yet for k-centers clustering and HC, the same number of clusters as in the integrated method is not necessarily the most suitable. Therefore, in **Figure 6**, the number of clusters for both k-centers clustering



and HC is adopted from 6 all the way to 100 for DS1 and DS2. Each number of clusters corresponds to 50 runs of k-centers clustering and a single run of HC. The numbers of clusters, ranging from 6 to 100, have corresponding *JCs*. The maximum *JC* for HC is identified and adopted. For k-centers clustering,

each corresponding number of clusters has to run 50 times, and there are consequently 50 *JCs*, whose mean is then calculated and acquired. The number of clusters ultimately gives rise to 95 mean values. We then adopt the number of clusters that corresponds with the biggest mean value, which gives a graphic representation

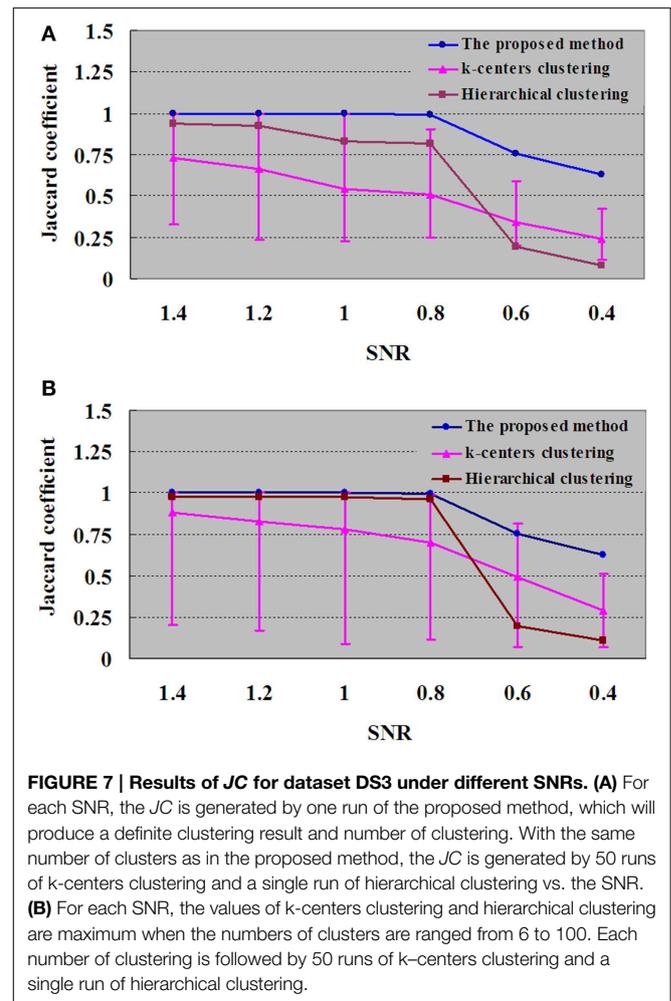


of the mean, maximum and minimum  $JC$  in relation to this number of cluster under k-centers clustering.

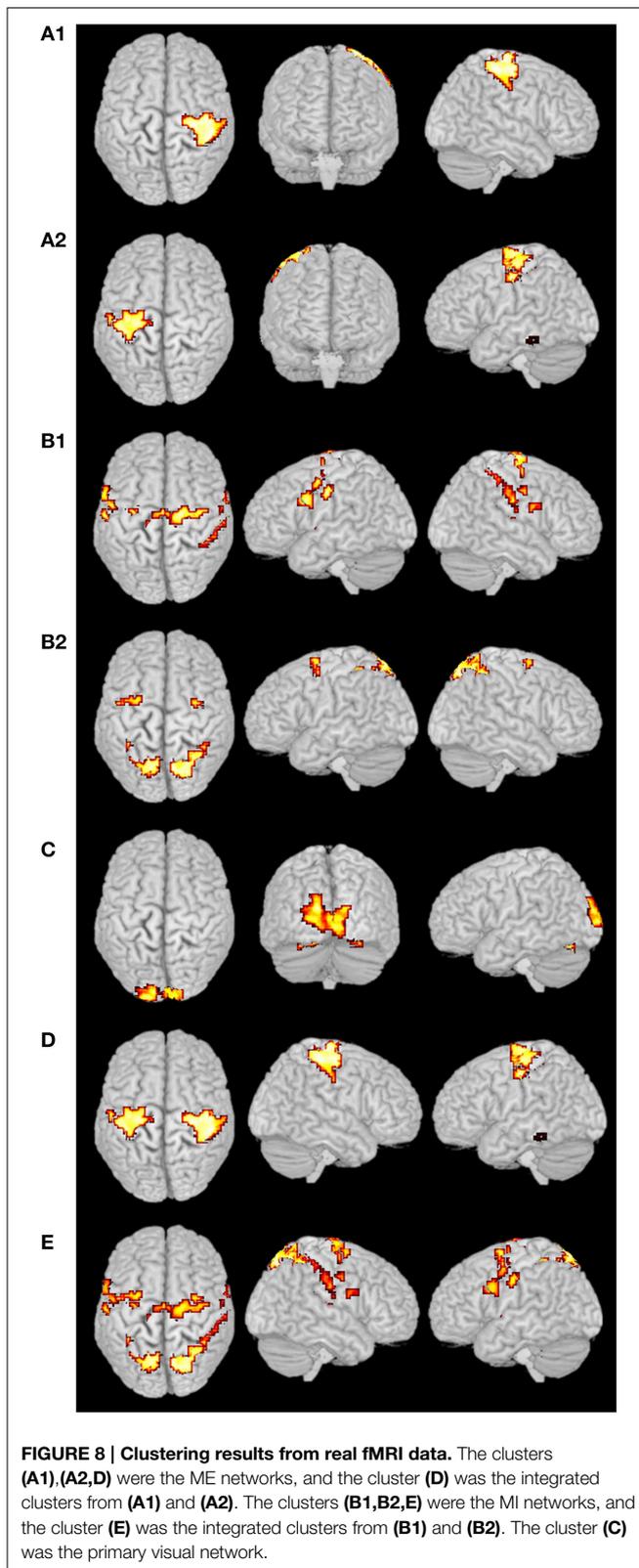
To test the performance of the algorithm with decreased SNR, the SNRs of DS3 were added, ranging from 1.4 to 0.4. **Figure 7** displays the  $JC$  curves obtained from the integrated method, k-centers clustering and HC with different SNRs.  $JC$  analysis is implemented to evaluate the performance of the proposed method and to compare with k-centers clustering and HC methods, and the results are given in **Figure 7**, in which the blue curve demonstrated that the performances turns to the downside when SNRs were varied from 0.8 to 0.4. In **Figure 7A**, for each SNR the  $JC$  is generated by one run of the proposed method, which will produce a definite clustering result and number of clusters. With the same number of clusters as in the proposed method, the  $JC$  is generated by 50 runs of k-centers clustering and a single run of HC vs. the SNR. In addition, in **Figure 7B**, for each SNR the number of clusters for k-centers clustering and HC are ranged from 6 to 100, and the values of k-centers clustering and HC are found to be maximum in the same range. And each number of clusters would give rise to 50 runs of k-center clustering and a single run of hierarchical clustering.

### fMRI Experimental Data Test

For the fMRI experiments in this study, subjects were presented with visual stimuli, bimanual MI and ME. These tasks had different durations and were performed sequentially. The primary clustering of the real fMRI data was completed by SOM, which produced relatively less amount of data than by PCA. Then we adopted the developed SOM-SAPC, which were able to reveal the new neural findings and identify the brain networks for both MI and ME using fMRI data. In particular, **Figure 8** shows the five clusters of activated cortical networks with the different stimulation patterns while **Figure 9** plots the calculated correlation coefficients between the mean time courses of the activation clusters and the different stimulation patterns of tasks in **Figures 8C–E**. Importantly, the time courses inside a specific cluster were found to be strongly correlated with a certain stimulus pattern. For example, a high correlation



coefficient (the coefficient value 0.817 is shown on the third row of **Table 2**) was revealed (shown in **Figure 9B** and **Table 2**) between the mean time courses of components in the clusters (**Figures 8A1,A2**) and the convoluted experiment design of the ME task. These findings suggested that the clusters were responsible for ME network (shown in **Figure 8D**, clusters (a1) and (a2) integrated). Moreover, the time courses of the clusters in **Figures 8B1,B2** were seen to be highly related to the convoluted stimulation pattern of the MI task (shown in **Figure 9A** and the coefficient value 0.622 is shown in the second row of **Table 2**), which indicated that these clusters should be related to MI network (shown in **Figure 8E**, clusters (b1) and (b2) integrated). Furthermore, the time course of the cluster in **Figure 8C** was found to be correlated with the convoluted stimulation pattern of the visual cue (shown in **Figure 9C** and **Table 2**). Finally, we observed that the spatial locations of the clusters were consistent with those captured by our previous studies (Wang et al., 2010) for the same datasets. In addition, it is very difficult to use these schemes (HC and k-centers clustering) to directly conduct pattern clustering due to the large datasets involved in the analysis of ME and MI functional networks. As such, only part of the datasets from the six slices of the whole-brain are adopted for



computation. The comparisons among the proposed SOM-SAPC scheme with other clustering approaches were also performed using the average squared error between the data points and the

center of the identified cluster (see **Figure 10**). We can see from **Figure 10** (the number of clusters is 66) that the error calculated from the SOM-SAPC was lower than that from the hierarchical clustering and the mean with 100 runs of *k*-centers clustering.

## Discussion

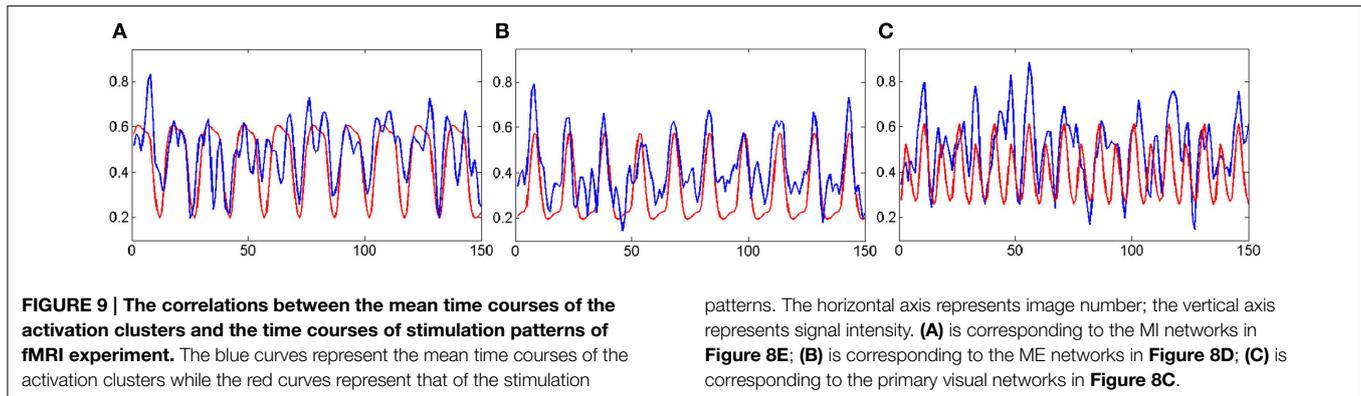
Recently identifying brain networks when brain is activated with multitask has attracted particular attention in the neurocognitive fields (Wang et al., 2010). For example, MI and ME have been extensively investigated using a number of multitask tests with fMRI measurements (Hanakawa et al., 2003, 2008; Hu et al., 2005; Wang et al., 2010). To classify the cerebral functions evoked in experiments, clustering methods were generally implemented to identify brain functional activation patterns (Chuang et al., 1999; Goutte et al., 1999; Ngan and Hu, 1999; Fadili et al., 2000; Mezer et al., 2009). However, few studies based on pattern clustering methods have been conducted using fMRI data for MI and ME network analysis. The PCA-SAPC proposed previously by our group (Zhang et al., 2011) also failed to effectively process the large datasets in order to classify and generate the ME and MI functional networks. In this study, we presented a multi-voxel clustering method, known briefly as SOM-SAPC, by integrating improved SOM and SAPC. This newly developed method is proposed to construct the cerebral networks for MI and ME, enabling the recognition of different signal patterns of brain activation by fMRI data with reduced computational cost. Our results suggested that brain networks and associated brain activation patterns were strongly correlated with different conditions of multitask experiments.

### Discussion on Conventional Clustering Methods

Several data-driven based clustering methods have been utilized to reveal brain activity regions with different signal patterns, (Barth et al., 1997; Baumgartner et al., 1998, 2000; Baune et al., 1999; Chuang et al., 1999; Filzmoser et al., 1999; Moser et al., 1999; Dimitriadou et al., 2004; Mezer et al., 2009), which includes *k*-center analysis (MacQueen, 1967) and hierarchical clustering (HC) analysis tools (Goutte et al., 1999; Chen et al., 2006; Liao et al., 2008; Ren et al., 2014). However, compared with SOM-SAPC, *k*-center analysis cannot resolve the problems like local minima and initialization when implemented with repeated runs. Further, unlike SOM-SAPC, stopping criteria has to be pre-determined for HC. Importantly, the data point pattern for HC can only be classified to a pre-denoted cluster, but not the other clusters (Jain and Dubes, 1988; Frigui and Krishnapuram, 1999; Geva, 1999; Berkhin, 2006).

### Discussion on Kohonen's SOM and PCA

Kohonen's SOM is considered an ingenious neural network which uses a simple geometric computation to replace the more detailed properties of the Hebb-like rule and lateral interactions (Haykin, 1999). For example, SOM has the disadvantages of using the Euclidean distance metric for fMRI signal analysis and for identifying the matched nodes. Compared with the correlation distance metric, Euclidean distance metric is more vulnerable to the baseline level. However, for MI and ME data processing, SOM



**TABLE 2 | The correlation coefficients calculated between the mean time courses of the activation clusters and the time courses of stimulation patterns of fMRI.**

| Cluster   | Correlation coefficient | Responding task |
|-----------|-------------------------|-----------------|
| Figure 8E | 0.622                   | Motor imagery   |
| Figure 8D | 0.817                   | motor execution |
| Figure 8C | 0.513                   | Visual cue      |

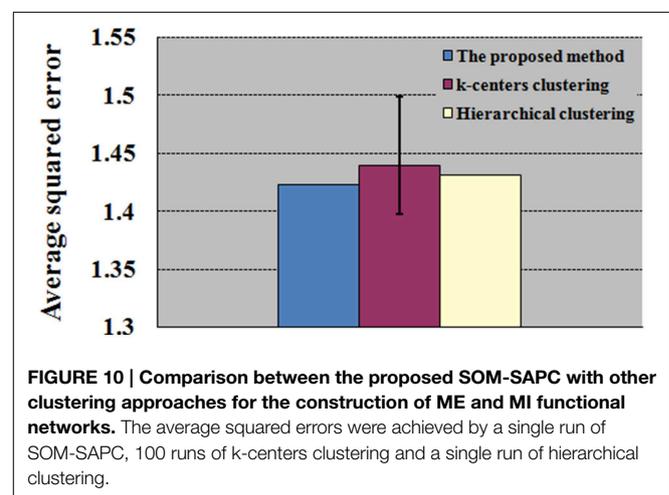
is still very effective in coping with the time course of whole-brain voxels, which can only produce hundreds of exemplar time courses before the later implementation of SAPC analysis. In contrast, although PCA can also be used to decrease the amount of data prior to SAPC clustering analysis, its mitigating effects on MI and ME data are far from satisfying, because it routinely fails to contribute to a smooth pattern clustering analysis when implemented with SAPC on PCs.

## Discussion on SAPC

SAPC is used to cluster the preliminary results generated by improved SOM, enabling a more complete and in-depth analysis of brain function. In SAPC, the Silhouette index is adopted to identifying the best clustering solution. Among many validity indices of clustering results, the Silhouette index performs very well in terms of calculating the number of clusters with clear cluster structures, which can also be used to judge the quality of the clustering result (Kaufman and Rousseeuw, 1990; Berkhin, 2006). In one-dimensional optimization of preference  $p$  under SAPC, we have adopted, unlike previously work (Zhang et al., 2011), the Golden-section search, so as to make sure of the accuracy of  $p$ -value.

## Discussion on the Integration of Methods

Compared with single only APC method, the developed SOM-SAPC approach is immune to possible memory resource insufficiency in analyzing large fMRI datasets, and able to significantly improve computation efficiency. The results in Figures 4B, 5B, 6, 7 show that the SOM-SAPC has better performance compared to those from  $k$ -centers clustering and HC in some respects. For example, the initial guesses of



exemplars is essential to  $k$ -center clustering method, in which a bunch of runs with different initial guesses are needed to get more accurate results based on this method (Frey and Dueck, 2007). In contrast, this is not the case for the integrated SOM-SAPC method since we can get better results by supervising the Silhouette index. As such, SOM-SAPC is different with the  $k$ -centers clustering because it is independent of the initialization of the cluster centers and further provides the optimal set of clusters for a given dataset.

## The Integration of Functional Networks

Figure 8 shows the overlapped activated regions in ME and MI during the associated tasks, which reveals clusters that represent the activation of different functional networks in task-related states. It should be pointed out that the findings in Figure 8 were in good agreement with the previously reported results (Lotze et al., 1999; Hanakawa et al., 2003; Lacourse et al., 2005; Wang et al., 2010). In our experiment, subjects were in a variety of task states, and each task state has its unique functional networks. Because of the functional asymmetry of cortical activity in motor areas during sequential finger movements (White et al., 1994; Amunts et al., 1996, 1997a,b, 2000; Dassonville et al., 1997; Volkman et al., 1998; Yuan and Ye, 2013; Pool et al., 2014), the left and right motor areas show

the differences in BOLD signal magnitude changes. Basically, activated regions responding to ME tasks were classified into two clusters (i.e., **Figures 8A1,A2**). For example, the activated regions in **Figures 8A1,A2** could be denoted as the same functional network (shown in **Figure 8D**) since the averaged time courses of components within these clusters were found to have strong correlation with the stimulation pattern of the ME task as compared to the corrections from other tasks (see **Figure 9B** and **Table 2**). The activated regions in **Figures 8B1,B2** were able to be classified into one MI networks (shown in **Figure 8E**) because we found the mean time courses from components inside these specific clusters (**Figures 8B1,B2**) were shown to be strongly correlated with the convoluted experiment design of the responding tasks (**Figure 9A** and **Table 2**). In addition, the clusters in **Figures 8B1,B2** consist of the same brain activity areas as previous findings (Hanakawa et al., 2003), which indicated brain networks integrated among the activated regions were essential to complete the MI task in our experiment.

## Conclusion

In this study, we proposed a novel method, namely SOM-SAPC, which is able to integrate improved SOM and SAPC for the clustering of brain activity patterns of multitasks related

to MI and ME. The developed method is able to handle large datasets using PC and is very efficient for brain connectivity analysis while the clustering method based on the combination of PCA and APC could hardly cope with such large datasets for construction of brain networks. Our fMRI findings validated that SOM-SAPC could reveal patterns of brain activity in response to multitasks. In particular, several unique brain networks that involved in performing the multitasks related to MI and ME were identified, which include parts of the visual, ME and MI functional networks. We suggested that the developed scheme is very sensitive to fMRI measurements and can effectively capture brain activity patterns and networks using fMRI data.

## Acknowledgments

This research was supported by the Natural Science Foundation of China (61273361, 61373009, and 61373091) and the Key Technology R&D Program of Sichuan Province (2012SZ0159, Science & Technology Department of Sichuan Province, China). The study was also supported by SRG2013-00035-FHS Grant, MYRG2014-00093-FHS Grant, MYRG 2015-00036-FHS from University of Macau in Macau and FDCT grant 026/2014/A1 from Macao government.

## References

- Amunts, K., Jäncke, L., Mohlberg, H., Steinmetz, H., and Zilles, K. (2000). Interhemispheric asymmetry of the human motor cortex related to handedness and gender. *Neuropsychologia* 38, 304–312. doi: 10.1016/S0028-3932(99)0075-5
- Amunts, K., Schlaug, G., Jäncke, L., Steinmetz, H., Schleicher, A., Dabringhaus, A., et al. (1997a). Motor cortex and hand motor skills: structural compliance in the human brain. *Hum. Brain Mapp.* 5, 206–215.
- Amunts, K., Schlaug, G., Schleicher, A., Steinmetz, H., Dabringhaus, A., Roland, P. E., et al. (1996). Asymmetry in the human motor cortex and handedness. *Neuroimage* 4, 216–222. doi: 10.1006/nimg.1996.0073
- Amunts, K., Schmidt-Passos, F., Schleicher, A., and Zilles, K. (1997b). Postnatal development of interhemispheric asymmetry in the cytoarchitecture of human area 4. *Anat. Embryol.* 196, 393–402.
- Anderberg, M. R. (1973). *Cluster Analysis for Applications*. New York, NY: Academic Press.
- Antoniu, A., and Lu, W. S. (2007). *Practical Optimization: Algorithms and Engineering Applications*. New York, NY: Springer, 92–95. doi: 10.1007/978-0-387-71107-2
- Barth, M., Diemling, M., and Moser, E. (1997). Modulation of signal changes in gradient-recalled echo functional MRI with increasing echo time correlate with model calculations. *Magn. Reson. Imag.* 15, 745–752. doi: 10.1016/S0730-725X(97)00043-X
- Baumgartner, R., Rynner, L., Richter, W., Summers, R., Jarmasz, M., and Somorjai, R. (2000). Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering vs. principal component analysis. *Magn. Reson. Imag.* 18, 89–94. doi: 10.1016/S0730-725X(99)00102-2
- Baumgartner, R. B., Windischberger, C., and Moser, E. (1998). Quantification in functional magnetic resonance imaging: fuzzy clustering vs. correlation analysis. *Magn. Reson. Imag.* 16, 115–125. doi: 10.1016/S0730-725X(97)00277-4
- Baune, A., Sommer, F. T., Erb, M., Wildgruber, D., Kardatzki, B., Palm, G., et al. (1999). Dynamical cluster analysis of cortical fMRI activation. *Neuroimage* 9, 477–489. doi: 10.1006/nimg.1999.0429
- Berkhin, P. (2006). *A Survey of Clustering Data Mining Techniques. Grouping Multidimensional Data: Recent Advances in Clustering*. New York, NY: Springer, 25–71. doi: 10.1007/3-540-28349-8\_2
- Brouziyne, M., and Molinaro, C. (2005). Mental imagery combined with physical practice of approach shots for golf beginners. *Percept. Mot. Skills* 101, 203–211. doi: 10.2466/pms.101.5.203-211
- Buxton, R. B., Uluda, K., Dubowitz, D. J., and Liu, T. T. (2004). Modeling the hemodynamic response to brain activation. *Neuroimage* 23, S220–S233. doi: 10.1016/j.neuroimage.2004.07.013
- Chen, H., Yuan, H., Yao, D., Chen, L., and Chen, W. (2006). An integrated neighborhood correlation and hierarchical clustering approach of functional MRI. *IEEE Trans. Biomed. Eng.* 53, 452–458. doi: 10.1109/TBME.2005.8696660
- Chen, H. F., Yang, Q., Liao, W., Gong, Q., and Shen, S. (2009). Evaluation of the effective connectivity of supplementary motor areas during motor imagery using Granger causality mapping. *Neuroimage* 47, 1844–1853. doi: 10.1016/j.neuroimage.2009.06.026
- Chuang, K. H., Chiu, M. J., Lin, C. C., and Chen, J. H. (1999). Model-Free functional MRI analysis using Kohonen clustering neural network and fuzzy C-means. *IEEE Trans. Med. Imag.* 18, 1117–1128. doi: 10.1109/42.819322
- Dassonville, P., Zhu, X. H., Ugurbil, K., Kim, S. G., and Ashe, J. (1997). Functional activation in motor cortex reflects the direction and the degree of handedness. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14015–14018. doi: 10.1073/pnas.94.25.14015
- Dimitriadou, E., Barth, M., Windischberger, C., Hornik, K., and Mose, E. (2004). A quantitative comparison of functional MRI cluster analysis. *Artif. Intell. Med.* 31, 57–71. doi: 10.1016/j.artmed.2004.01.010
- Ding, M., Chen, Y., and Bressler, S. L. (2006). “Granger causality: basic theory and application to neuroscience,” in *Handbook of Time Series Analysis*, eds B. Schelter, M. Winterhalder, and J. Timmer (Weinheim: Wiley-VCH Verlag), 437–460.
- Fadili, M. J., Ruan, S., Bloyet, D., and Mazoyer, B. (2000). A multistep unsupervised fuzzy clustering analysis of fMRI time series. *Hum. Brain Mapp.* 10, 160–178. doi: 10.1002/1097-0193(200008)10:4<160::AID-HBM20>3.0.CO;2-U
- Felton, E. A., Wilson, J. A., Williams, J. C., and Garell, P. C. (2007). Electrocorticographically controlled brain-computer interfaces using motor and sensory imagery in patients with temporary subdural electrode implants. Report of four cases. *J. Neurosurg.* 106, 495–500. doi: 10.3171/jns.2007.106.3.495

- Filzmoser, P., Baumgartner, R., and Moser, E. (1999). A hierarchical clustering method for analyzing functional MR images. *Magn. Reson. Imag.* 17, 817–826. doi: 10.1016/S0730-725X(99)00014-4
- Frey, B. J., and Dueck, D. (2007). Clustering by passing messages between data points. *Science* 315, 972–976. doi: 10.1126/science.1136800
- Frigui, H., and Krishnapuram, R. (1999). A robust competitive clustering algorithm with applications in computer vision. *IEEE Trans. Pattern Anal. Mach. Intell.* 21, 450–465. doi: 10.1109/34.765656
- Geva, A. B. (1999). Hierarchical unsupervised fuzzy clustering. *IEEE Trans. Fuzzy Syst.* 7, 723–733. doi: 10.1109/91.811242
- Goutte, C., Toft, P., Rostrup, E., Nielsen, F. Å., and Hansen, L. K. (1999). On clustering fMRI time series. *Neuroimage* 9, 298–310. doi: 10.1006/nimg.1998.0391
- Guan, R., Shi, X., Marchese, M., Yang, C., and Liang, Y. (2011). Text clustering with seeds affinity propagation. *IEEE Trans. Knowl. Data Eng.* 23, 627–637. doi: 10.1109/TKDE.2010.144
- Hanakawa, T., Dimyan, M. A., and Hallett, M. (2008). Motor planning, imagery, and execution in the distributed motor network distributed motor network: a time-course study with functional MRI. *Cereb. Cortex* 18, 2775–2788. doi: 10.1093/cercor/bhn036
- Hanakawa, T., Immisch, I., Toma, K., Dimyan, M. A., Gelderen, P. V., and Hallett, M. (2003). Functional properties of brain areas associated with motor execution and imagery. *J. Neurophysiol.* 89, 989–1002. doi: 10.1152/jn.00132.2002
- Hausfeld, L., Valente, G., and Formisano, E. (2014). Multiclass fMRI data decoding and visualization using supervised self-organizing maps. *Neuroimage* 96, 54–66. doi: 10.1016/j.neuroimage.2014.02.006
- Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation, 2nd Edn.* Englewood Cliffs, NJ: Prentice Hall, 465–500.
- Hu, D. W., Yan, L. R., Liu, Y. D., Zhou, Z. T., Friston, K. J., Tan, C. L., et al. (2005). Unified SPM-ICA for fMRI analysis. *Neuroimage* 25, 746–755. doi: 10.1016/j.neuroimage.2004.12.031
- Jain, A. K., and Dubes, R. C. (1988). *Algorithms for Clustering Data.* Englewood Cliffs, NJ: Prentice Hall.
- Jeanerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia* 33, 1419–1432. doi: 10.1016/0028-3932(95)00073-C
- Kaufman, L., and Rousseeuw, P. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis.* London: John Wiley and Sons. ISBN: 10: 0471878766.
- Kimberley, T. J., Khandekar, G., Skrabala, L. L., Spencer, J. A., Van Gorp, E. A., and Walker, S. R. (2006). Neural substrates for motor imagery in severe hemiparesis. *Neurorehabil. Neural Repair* 20, 268–277. doi: 10.1177/1545968306286958
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biol. Cybernet.* 43, 59–69. doi: 10.1007/bf00337288
- Kohonen, T. (1990). The self-organizing map. *Proc. IEEE* 78, 1464–1480. doi: 10.1109/5.58325
- Kohonen, T. (1995). *Self-organizing Maps.* New York, NY: Springer Verlag, 362.
- Lacourse, M. G., Orr, E. L., Cramer, S. C., and Cohen, M. J. (2005). Brain activation during execution and motor imagery of novel and skilled sequential hand movements. *Neuroimage* 27, 505–519. doi: 10.1016/j.neuroimage.2005.04.025
- Liao, W., Chen, H. F., Yang, Q., and Lei, X. (2008). Analysis of fMRI data using improved self-organizing mapping and spatio-temporal metric hierarchical clustering. *IEEE Trans. Med. Imag.* 27, 1472–1483. doi: 10.1109/TMI.2008.923987
- Lotze, M., and Halsband, U. (2006). Motor imagery. *J. Phys. Paris* 99, 386–395. doi: 10.1016/j.jphysparis.2006.03.012
- Lotze, M., Montoya, P., Erb, M., Hulsmann, E., Flor, H., Klose, U., et al. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *J. Cogn. Neurosci.* 11, 491–501. doi: 10.1162/089892999563553
- MacQueen, J. (1967). “Some methods for classification and analysis of multivariate observations,” *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley, CA: University of California Press), 281–297.
- Malouin, F., Richards, C. L., Doyon, J., Desrosiers, J., and Belleville, S. (2004). Training mobility tasks after stroke with combined mental and physical practice: a feasibility study. *Neurorehabil. Neural Repair* 18, 66–75. doi: 10.1177/0888439004266304
- Mézard, M. (2007). Where are the good exemplars?. *Science* 315, 949–951. doi: 10.1126/science.1139678
- Mezer, A., Yovel, Y., Pasternak, O., Gorfine, T., and Assaf, Y. (2009). Cluster analysis of resting-state fMRI time series. *Neuroimage* 45, 1117–1125. doi: 10.1016/j.neuroimage.2008.12.015
- Moser, E., Baumgartner, R., Barth, M., and Windischberger, C. (1999). Explorative signal processing in functional MR imaging. *Int. J. Imag. Syst. Technol.* 10, 166–176.
- Ngan, S. C., and Hu, X. (1999). Analysis of functional magnetic resonance imaging data using self-organizing mapping with spatial connectivity. *Magn. Res. Med.* 41, 939–946.
- Ngan, S. C., Yacoub, E. S., Auffermann, W. F., and Hu, X. (2002). Node merging in Kohonen's self-organizing mapping of fMRI data. *Artif. Intell. Med.* 25, 19–33. doi: 10.1016/S0933-3657(02)00006-4
- Peltier, S. J., Polk, T. A., and Noll, D. C. (2003). Detecting low-frequency functional connectivity in fMRI using a self-organizing map (SOM) algorithm. *Hum. Brain Mapp.* 20, 220–226. doi: 10.1002/hbm.10144
- Pfurtscheller, G., and Neuper, C. (2001). Motor imagery and direct brain-computer communication. *Proc. IEEE* 89, 1123–1134. doi: 10.1109/5.939829
- Pool, E. M., Rehme, A. K., Fink, G. R., Eickhoff, S. B., and Grefkes, C. (2014). Handedness and effective connectivity of the motor system. *Neuroimage* 99, 451–460. doi: 10.1016/j.neuroimage.2014.05.048
- Porro, C. A., Francescato, M. P., Cettolo, V., Diamond, M. E., Baraldi, P., Zuiani, C. H., et al. (1996). Primary motor and sensory cortex activation during motor performance and motor imagery: a functional magnetic resonance imaging study. *J. Neurosci.* 16, 7688–7698.
- Ren, T., Zeng, W., Wang, N., Chen, L., and Wang, C. (2014). A novel approach for fMRI data analysis based on the combination of sparse approximation and affinity propagation clustering. *Magn. Reson. Imag.* 32, 736–746. doi: 10.1016/j.mri.2014.02.023
- Volkman, J., Schnitzler, A., Witte, O. W., and Freund, H. J. (1998). Handedness and asymmetry of hand representation in human motor cortex. *J. Neurophysiol.* 79, 2149–2154.
- Wang, Y. Q., Chen, H. F., Gong, Q. Y., Shen, S., and Gao, Q. (2010). Analysis of functional networks involved in motor execution and motor imagery using combined hierarchical clustering analysis and independent component analysis. *Magn. Reson. Imag.* 28, 653–660. doi: 10.1016/j.mri.2010.02.008
- White, L. E., Lucas, G., Richards, A., and Purves, D. (1994). Cerebral asymmetry and handedness. *Nature* 368, 197–198. doi: 10.1038/368197a0
- Yuan, Z., and Ye, J. (2013). Fusion of fNIRS and fMRI: identification when and where hemodynamic responses are changing in human brain. *Front. Hum. Neurosci.* 7:676. doi: 10.3389/fnhum.2013.00676
- Zhang, J., Tuo, X. G., Yuan, Z., Liao, W., and Chen, H. F. (2011). Analysis of fMRI data using an integrated principal component analysis and supervised affinity propagation clustering approach. *IEEE Trans. Biomed. Eng.* 58, 3184–3196. doi: 10.1109/TBME.2011.2165542

**Conflict of Interest Statement:** 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 © 2015 Zhang, Liu, Chen, Yuan, Huang, Deng, Lu, Zhang, Wang, Wang and Chen. 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) or licensor 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.