Spectral Clustering Approach with Sparsifying Technique for Functional Connectivity Detection in the Resting Brain

Mahdi Ramezani[1], Amin Heidari[1], Emad Fatemizadeh[1] and Hamid Soltanian-Zadeh[2, 3]
[1] Biomedical Signal and Image Processing Laboratory, Sharif University of Technology, Tehran, Iran
[2] Radiology Image Analysis Laboratory, Henry Ford Health System, Detroit, MI 48202 USA
[3] Control and Intelligent Processing Center of Excellence, Electrical and Computer Engineering Department, University of Tehran, Iran

Abstract- **The aim of this study is to assess the functional connectivity from resting state functional magnetic resonance imaging (fMRI) data. Spectral clustering algorithm was applied to the realistic and real fMRI data acquired from a resting healthy subject to find functionally connected brain regions. In order to make computation of the spectral decompositions of the entire brain volume feasible, the similarity matrix has been sparsified with the t-nearest-neighbor approach. Realistic data were created to investigate the performance of the proposed algorithm and comparing it to the recently proposed spectral clustering algorithm with the Nystrom approximation and also with some well-known algorithms such as the Cross Correlation Analysis (CCA) and the spatial Independent Component Analysis (sICA). To enhance the performance of the methods, a variety of data pre and post processing steps, including data normalization, outlier removal, dimensionality reduction by using wavelet coefficients, estimation of number of clusters and optimal number of independent components (ICs). Results demonstrate the applicability of the proposed algorithm for functional connectivity analysis.**

I. INTRODUCTION

Functional neuroimaging has become an important neuroscientific tool for assessing connectivity and interactions of distant areas of the brain. Scientists have defined two types of connectivity: effective connectivity which indicates the influence that a neural system exerts over another and functional connectivity defined as the temporal correlations between spatially remote neurophysiological events [1].

Various modalities like electroencephalographic (EEG), magnetoencephalographic (MEG), positron emission tomography (PET) and functional resonance imaging (fMRI) may be used to determine regional interactions in the brain. Among them, fMRI as a non-invasive imaging modality with high spatial resolution has drawn considerable attention. The fMRI data in the resting state (while subjects perform no task) is widely used for exploring functional connectivity. During the resting state, there are intrinsic hemodynamic fluctuations in the blood oxygen level dependent (BOLD) signal. The resting state fMRI is used to analyze the BOLD signal using the similarity of fluctuations in different voxels. Functional connectivity is evaluated by cross correlating temporal BOLD signals between brain regions [2]. Usually a region of interest (ROI) is selected in the brain and correlations between the averaged time course from voxels of this region (which is used as a reference signal) and all other voxels are computed. These kinds of approaches are called model-based methods, because they are based on prior knowledge. Alternatives to them are data-driven methods that need no predefined region [3].

Independent Component Analysis (ICA) is a widely used data-driven method that needs no prior knowledge about the spatial or temporal patterns of source signals. As a result, ICA is an appropriate method for studying the resting state data, which is increasingly being used. Although ICA was introduced into fMRI data processing in 1998 [4], its first application for analyzing resting state fMRI data was reported in 2003 [5]. A problem with ICA is the choice of meaningful decomposed components.

Another class of data-driven methods that have become popular are clustering methods. Among the methods in this class are K-means clustering, fuzzy c-means clustering, and hierarchical clustering. Like almost any other clustering method, the K-means results depend on the assignment of the initial clusters. Moreover, the number of clusters must be specified beforehand [6]. In 1998, Golay et al. employed FCM clustering with the inverse of the similarity between time series as a distance measure for functional connectivity detection [7]. A limitation of using FCM is the need to choose the number of clusters in advance. Besides the structured noises can contaminate the proposed distance metrics and degrade the performance. In order to overcome these problems, Cordes et al applied a hierarchical clustering algorithm based on single linkage (nearest neighbour method) to measure the connectivity in the fMRI resting state data [8]. In this method, each voxel is considered as a cluster at the beginning and close clusters are merged to create bigger clusters through the linkage function in the next stages. The algorithm chooses the number of clusters by selecting a linkage inconsistency threshold and partitions the data. However, the resulting partitions are not as homogeneous as with the K-means. In addition, because of the increasing constraints on the partition at each stage, partitioning the data to a small number of clusters can result in partitions considerably worse than partitions obtained with the K-means [6]. In 2008, using the K-means clustering and short time frequency analysis, Mezer et al. studied the spatial signal characteristics of the resting state fMRI time series [9].

As well, all of the mentioned clustering methods assume that data clusters conform to certain shapes, while on the contrary this is not a valid assumption for the fMRI data. A successful clustering method that does not make any assumption on the form of the clusters, is spectral clustering [10] which can solve all of the above problems. This method relies on the eigen decomposition of the similarity matrix between pairs of voxels.

Spectral clustering is introduced into the fMRI analysis in 2005 by Lin et al. for activation detection [11]. In 2009, Venkataraman et al. applied it for analyzing the resting state fMRI data [12]. Due to the large number of voxels, they used the Nystrom approximation to estimate the eigenvalues and eigenvectors.

In this paper, we apply spectral clustering with the strategy of sparsifying the similarity matrix via keeping nearest neighbours. So, minor connections between data points are neglected and some noisy/inaccurate similarity values might be discarded. We have applied the method to the realistic and experimental fMRI. Realistic data were used to investigate the influence of the noise level (contrast to noise ratio) on the performances of the algorithm. Also, variety of data pre and post processing steps were employed. The proposed approach was compared to the recently developed spectral clustering algorithm based on nystrom approximation [12], Cross Correlation Analysis and two commonly used ICA algorithm, the infomax [13] and the fixed-point[14]. To the best of our knowledge, this is one of the first of its kind to use the t-nearest neighbor approach [15] for the analysis of the resting state fMRI. Besides, comparing the results of spectral clustering and ICA approaches, on the resting state brain data, have not been performed in the past.

II. MATERIALS

A. Experimental Data

A dataset provided by the Pittsburgh Brain Competition (PBC) (http://www.braincompetition.org) was used in this study. Data was collected on a subject (male, 27 years) on a Siemens 3T TIM Trio with 32 channel head coil at the Magnetic Resonance Research Center of the University of Pittsburgh Medical Center in compliance with their Internal Review Board regulations. Each acquisition consisted of 50 contiguous slices (96×96×50 2.4mm × 2.4mm × 2.4mm voxels). The repetition time (TR) was set arbitrarily to 2s. At whole 300 acquisitions were made from the subject.

We performed realignment using SPM8 (http://www.fil.ion.bpmf.ac.uk/spm/). The data linear trend was removed and temporally filtered using a bandpass filter with 0.01Hz and 0.1Hz cut-offs using RESTing-state fMRI data analysis toolkit (http://resting-fmri.sourceforge.net).

The brain scoring was also prepared by the 2009 Organization of Human Brain Mapping. Each scoring map contains one functional region that was computed from fMRI image data and was the target region to be found (shown with “white” voxels in Figure 1) and a number of voxels very near to this region that were considered too close to call for inclusion in the functional region (“gray” voxels). Participants in the competition would be neither rewarded nor penalized for including these voxels in a submitted region.

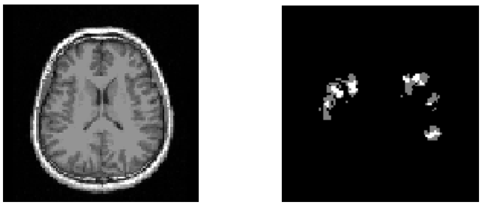
(a) (b)

Figure 1. Slice number 26: (a) T_1-weigthed anatomical image, (b) Scoring map.

B. Realistic Data

For creating the realistic data we have used the experimental data which was described in previous subsection. After realigning the data set using SPM8, the time course of each voxel was randomized to remove the intrinsic connectivity in the whole dataset. The 19 non-overlapping functionally connected regions were chose from the brain scoring maps provided from the PBC 2009. For each region a connectivity signal was created and added to the time courses of voxels of each region. The amplitude of the created signal was chose to be 4% of the mean amplitude of the time course of the selected voxel. Then the data linear trend was removed and temporally filtered using a bandpass filter with 0.01Hz and 0.1Hz cut-offs using RESTing-state fMRI data analysis toolkit.

The synthetic connectivity signals were generated as below. A random superposition of sines and cosines oscillating at various frequencies can be used to represent any stationary time-series [16]. Therefore, a linear combination of sine waveforms of different frequencies and phases (weights of each generated from a normal distribution) were used to create each regional time-series. In order to produce a temporally smoothed BOLD time-series, the regional time-series were then convolved with a hemodynamic response function (double gamma). Synthetic connectivity between regional time-series was achieved by linking the weights of sine waveforms at certain frequencies and phases [17].

C. Preprocessings

In order to enhance the performance, a variety of data preprocessing steps were employed. Two of them were done for all methods. First, the data were normalized by subtracting the mean value and dividing by the overall standard deviation. Thus, each voxel had mean activity of zero and unit standard deviation. Secondly, the outliers were removed by setting all values that were beyond 3 standard deviations from the mean to a fixed value of 3 or −3, depending on the sign of the original value.

III. METHODS

A. Spectral Clustering

Our spectral clustering method requires the data to be in matrix form. The colums of the data matrix (data points $x_1,\ldots,x_n$) contain the signal intensity measured at each voxel at each time point. So the data matrix is a t×v matrix (t = number of scans, v = number of voxels). The spectral clustering uses the similarity matrix $S=(s_{ij})_{i,j=1..n}$ to group the data points into k clusters. The s_{ij} is nonnegative and symmetric and reflects the relationship between x_i and x_j. In this study, we have used a normalized spectral clustering, a common variant of the spectral clustering as follows [10]:

1) Use Gaussian as a similarity function.

$$S_{ij} = \exp(-\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}) \tag{1}$$

2) Compute the normalized Laplacian matrix

$$L = I - D^{-1/2} S D^{-1/2} \tag{2}$$

where D is a diagonal matrix with

$$D_{ii} = \sum_{j=1}^{n} S_{ij} \; .$$

3) Compute the first k eigenvectors $v_1, \ldots, v_k$ of L.
4) Create the matrix V containing the vectors $v_1, \ldots, v_k$ as columns ($V \in R^{n \times k}$).
5) Normalizing the row sums of V to have norm 1, that is $u_{ij} = v_{ij} \Big/ (\sum_k v_{ik}^2)^{1/2}$ ($U \in R^{n \times k}$).

Cluster the rows of U ($y_i \in R^k)_{i=1,\ldots,n}$ with the k-means algorithm into clusters $C_1, \ldots, C_k$.

An unpleasant aspect of spectral clustering is the large amount of memory used for storing S and the computational cost for eigen-decomposition of the matrix L, which is an N×N matrix where N is on the order of 200,000 voxels [15]. In order to solve this problem, Venkataraman et al. approximate the eigenvalues and eigenvectors of L using Nystrom approximation [12]. In this paper, we use sparsification method (t-Nearest-Neighbour (tNN) approach [15]) to reduce the memory use. So, the memory requirement reduces from $O(n^2)$ to O(nt). It is also possible to use sparse eigen-solver for the sparse similarity matrix. The matrix S is sparsified by retaining only the nearest neighbors. In other words, only sij where j(or i) is among the t nearest neighbors of i(or j) are kept and other values of sij are made zero. For each voxel, the Euclidean distances to all voxels are computed, as in (3) and the t-nearest-neighbors are found.

$$\left\|x_i - x_j\right\|^2 = \left\|x_i\right\|^2 + \left\|x_j\right\|^2 - 2x_i^T x_j \tag{3}$$

To conserve time, all $\left\|x_j\right\|^2$ of data are pre-computed. To symmetrize the sparse distance matrix, the (i,j) and (j,i) elements of the matrix are set to the same value.

For spectral clustering analysis, several choices should be made. First of all, the parameter t in tNN should be carefully set to govern the connectedness of the similarity graph. Generally, the tNN graph connects points on different scales. In order to use this helpful property, we have run the algorithm for our realistic data and chose the number 50. Secondly, the number of clusters, k, should be chose. Eigen gap heuristic is a tool which is designed to solve this issue [10]. The goal of this method is to find the relatively small eigenvalue λ_{k+1} and set the number of clusters to k. However, this method returns ambiguous results in cases that the clusters are noisy and overlapping. Thirdly, because of the high dimensionality of the data (fMRI time series), direct clustering of that may cause problems. So we have used the wavelet coefficients, extracted using discrete wavelet transform as the feature to be clustered. For each time series, we have computed the detail coefficients of the wavelet decomposition at level 3 using the Daubechies wavelet of order 4. We have performed spectral clustering analysis on both the wavelet coefficients and raw fMRI time series.

Finally, in order to obtain the best results, we have performed k-means algorithm five times using different random initializations.

B. *Independent Component Analysis*

Independent component analysis finds functionally connected patterns of activity that are as statistically independent as possible. The observed data X (the t×v matrix) is modelled as[18]:

$$X = MS \tag{4}$$

where M is the t×m mixing matrix and S is the m×v matrix whose rows indicate the spatially independent components (t = number of scans, v = number of voxels, m = number of components). Thus, the independent components can be found by the following linear transformation:

$$S = WX \tag{5}$$

where W is the pseudo inverse of M.

In order to solve the equation, we used two common algorithms: the Infomax and the fixed-point, both of which minimize the mutual information of the components. The Infomax adaptively maximizes the output entropy of a neural network and the Fixed-point algorithm (FastICA) uses a robust approximation of the negentropy as a contrast function and iteratively maximizes it.

After the ICA decomposition, spatial maps were Z-transformed to absolute value and sign and were thresholded. Only voxels with intensities above a threshold were considered to be a part of the functional connectivity network.

C. *Cross Correlation Analysis*

One of the most used resting-state fMRI processing methods is cross correlation analysis that needs a reference signal. Wernicke region was selected as a region of seeds for the realistic and experimental data. The averaged time courses from voxels of these regions were computed as a reference signal. Correlations between the reference signal and all other voxels were computed. Correlation map were then Z-transformed and thresholded.

IV. RESULTS

In this section the results for the simulated and experimental datasets are presented.

A. *Realistic Data*

The number of clusters was estimated by the eigen gap heuristic approach explained in previous section to be 4 for spectral clustering analysis.

Figure 2 shows the result of the functional connectivity analysis in the resting state realistic data.

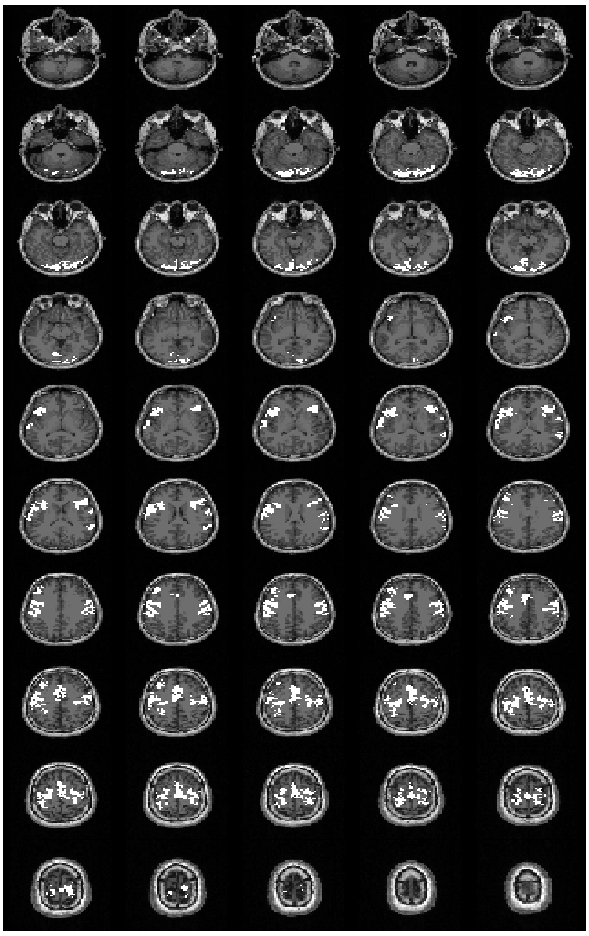

Figure 2. Functional connectivity networks detected by: SC-tNN.

In order to compare the results quantitatively, the Jaccard, Accuracy and the score coefficients were used as in (7).

$$J = \frac{TP}{TP + FN + FP}$$
$$ACC = \frac{TP + TN}{TP + TN + FN + FP} \tag{7}$$
$$Score = \frac{TP - FP}{TP}$$

where TP is the true positive, TN is the true negative, FP is the false positive and FN is the false negative. Table I shows the numerical results for the applied methods.

TABLE I
JACCARD, ACCURACY AND SCORE COEFFICIENTS FOR REALISTIC DATA

Method	Jaccard (%)	Accuracy (%)	Score (%)
SC-tNN	84.5	99.7	99.7
SC_tNN (Wavelet Coefficient)	48.8	98.9	100
SC-Nystrom	26.8	95.1	-156.6
CCA	63.4	99.0	67.1
FastICA	0.3	97.4	-6657
InfomaxICA	16.5	98.2	87.5

The results confirm that the proposed method can detect the connected regions without any a priori knowledge about the location of seed regions. It is clear from Table I that all algorithms can detect the connected regions with high accuracy rates which describes the high rate of correctly detected unconnected voxels. The proposed approach provides higher Jaccard coefficient rate which stands for greater number of correctly detected connected voxels and lower incorrectly detected connected and unconnected voxels. By using wavelet coefficients the dimension of the data and so the complexity of the clustering analysis is reduced without great change in the score rate but some of the connected voxels were not detected (FNs).

The spectral clustering with Nystrom approximation have large incorrectly detected voxels; therefore it is not suitable for functional connectivity detection. The reason is that the small similarity values are thrown aside in the tNN approach, so minor connections between data points are neglected. Therefore, by using the tNN approach not only little information is lost, but also some noisy/inaccurate similarity values might be discarded. On the other hand, there is no evidence that Nystrom approximation may provide higher quality results than using the fully dense matrix.

The CCA can detect the connected regions with an acceptable results but the requirement for prior knowledge constrains the exploration of possible functional connectivity and renders the detected functional connectivity sensitive to seed selection.

Table I also shows that the InfomaxICA algorithm has outperformed the FastICA algorithm which is similar to our previous study [21]. However, the results of spectral clustering analysis with tNN approach is superior to that.

B. *Experimental Data*

For the experimental data the eigengap heuristic approach fails to find the number of clusters (the differences between consequent eigenvalues were small). So the eigen gaps have been computed and compared to a threshold. The number of clusters has been set to the number of values above the threshold. This number was estimated to be 19. The number of sources estimated by the AIC estimation was 79 and the result of MDL was 33. So 56 was chosen as an estimate of the number of ICs.

The numerical result of the functional connectivity analysis in the experimental data is shown in Table II.

TABLE II
JACCARD, ACCURACY AND SCORE COEFFICIENTS FOR REAL DATA

Method	Jaccard (%)	Accuracy (%)
SC-tNN	9.5	98.7
SC_tNN (Wavelet Coefficient)	4.7	93.4
SC-Nystrom	6.7	98.4
CCA	1.1	96.8
FastICA	0.8	98.5
InfomaxICA	2.0	97.3

These results also confirm that the proposed method can detect the connected regions better than the other methods. However, the false positive rate is still high in the used experimental data.

V. CONCLUSION

We have addressed the problem of functional connectivity analysis in this paper. Spectral clustering algorithm with the sparsifying strategy was employed. Some of the existing data-driven methods have been reviewed and compared to the proposed algorithm. The experimental results demonstrate that the proposed method has solved the problem of eigen decomposition of the entire volume of the brain and provides a promising approach to find the connected regions.

REFERENCES

[1] K. Friston, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*: Academic Press, 2006.

[2] B. P. Rogers, V. L. Morgan, A. T. Newton *et al.*, "Assessing functional connectivity in the human brain by fMRI," *Magnetic Resonance Imaging,* vol. 25, no. 10, pp. 1347-1357, 2007.

[3] K. Li, L. Guo, J. Nie *et al.*, "Review of methods for functional brain connectivity detection using fMRI," *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society,* vol. 33, no. 2, pp. 131-139, 2009.

[4] M. McKeown, S. Makeig, G. Brown *et al.*, "Analysis of fMRI data by blind separation into independent spatial components," *Human Brain Mapping,* vol. 6, no. 3, pp. 160-188, 1998.

[5] V. Kiviniemi, J. H. Kantola, J. Jauhiainen *et al.*, "Independent component analysis of nondeterministic fMRI signal sources," *Neuroimage,* vol. 19, no. 2 Pt 1, pp. 253-260, 2003.

[6] C. Goutte, P. Toft, E. Rostrup *et al.*, "On clustering fMRI time series," *Neuroimage,* vol. 9, no. 3, pp. 298-310, 1999.

[7] X. Golay, S. Kollias, G. Stoll *et al.*, "A new correlation-based fuzzy logic clustering algorithm for fMRI," *Magn Reson Med,* vol. 40, no. 2, pp. 249-260, 1998.

[8] D. Cordes, V. Haughton, J. D. Carew *et al.*, "Hierarchical clustering to measure connectivity in fMRI resting-state data," *Magn Reson Imaging,* vol. 20, no. 4, pp. 305-317, 2002.

[9] A. Mezer, Y. Yovel, O. Pasternak *et al.*, "Cluster analysis of resting-state fMRI time series," *Neuroimage,* vol. 45, no. 4, pp. 1117-1125, 2008.

[10] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing,* vol. 17, no. 4, pp. 395-416, 2007.

[11] S. Lin, H. Pheng Ann, and W. Tien-Tsin, "A Spectral Clustering Approach to fMRI Activation Detection." pp. 5892-5895.

[12] A. Venkataraman, K. R. A. V. Dijk, R. L. Buckner *et al.*, "Exploring functional connectivity in fMRI via clustering," in Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing - Volume 00, 2009.

[13] A. Bell, and T. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput.,* vol. 7, no. 6, pp. 1129-1159, 1995.

[14] A. Hyvarinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis," *IEEE Transactions on Neural Networks,* vol. 10, no. 3, pp. 626-634, 1999.

[15] Y. Song, W.-Y. Chen, H. Bai *et al.*, "Parallel Spectral Clustering," *Machine Learning and Knowledge Discovery in Databases*, pp. 374-389, 2008.

[16] R. Shumway, and D. Stoffer, *Time Series Analysis and its Applications: With R Examples*: Springer, 2006.

[17] D. Zhou, W. K. Thompson, and G. Siegle, "MATLAB toolbox for functional connectivity," *NeuroImage,* vol. 47, pp. 1590–1607, 2009.

[18] V. G. van de Ven, E. Formisano, D. Prvulovic *et al.*, "Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest," *Hum Brain Mapp,* vol. 22, no. 3, pp. 165-178, 2004.

[19] Y.-O. Li, T. l. Adali, and V. Calhoun, "Estimating the number of independent components for functional magnetic resonance imaging data," *Hum Brain Mapp*, pp. 1251-1266, 2007.

[20] M. Ramezani, and E. Fatemizadeh, "Detecting Functional Connectivity in the Resting Brain using Independent Component Analysis," in 16th Iranian Conference on Biomedical Engineering, Tehran, Iran, 2009.