Brain Topogr (2009) 22:134–144
DOI 10.1007/s10548-009-0095-4

ORIGINAL PAPER

Detecting Functional Connectivity in fMRI Using PCA and Regression Analysis

Yuan Zhong · Huinan Wang · Guangming Lu · Zhiqiang Zhang · Qing Jiao · Yijun Liu

Received: 28 November 2008 / Accepted: 14 April 2009 / Published online: 1 May 2009

Abstract A fMRI connectivity analysis approach combining principal component analysis (PCA) and regression analysis is proposed to detect functional connectivity between the brain regions. By first using PCA to identify clusters within the vectors of fMRI time series, more energy and information features in the signal can be maintained than using averaged values from brain regions of interest. Then, regression analysis can be applied to the extracted principal components in order to further investigate functional connectivity. Finally, *t*-test is applied and the patterns with *t*-values lager than a threshold are considered as functional connectivity mappings. The validity and reliability of the presented method were demonstrated with both simulated data and human fMRI data obtained during behavioral task and resting state. Compared to the conventional functional connectivity methods such as average signal based correlation analysis, independent component analysis (ICA) and PCA, the proposed method achieves competitive performance with greater accuracy and true positive rate (TPR). Furthermore, the 'default mode' and motor network results of resting-state fMRI data indicate that using PCA may improve upon application of existing regression analysis methods in study of human brain functional connectivity.

Keywords Functional connectivity · Principal component analysis · Regression analysis · fMRI · Resting-state

Y. Zhong · H. Wang
Department of Biomedical Engineering, College of Automation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Y. Zhong · G. Lu (✉) · Z. Zhang · Q. Jiao
Department of Medical Imaging, Nanjing Jinling Hospital, Clinical School of Medical College, Nanjing University, 305 Eastern Zhongshan Road, Nanjing 210002, China
e-mail: cjr.luguangming@vip.163.com

Y. Liu
Departments of Psychiatry and Neuroscience, McKnight Brain Institute, University of Florida, Gainesville, Florida 32610, USA

Introduction

The concept of functional connectivity has been introduced into the functional neuroimaging literature as a descriptive measure of spatiotemporal correlations between spatially distinct regions of cerebral cortex (Friston et al. 1993; Lowe et al. 1998; Strother et al. 1995). Most resting-state functional magnetic resonance imaging (fMRI) processings performed so far used cross-correlation analysis (CCA), in which neuronal activations and functional connectivity detected by evaluating the correlation between the time course of each voxel and a reference function (Biswal et al. 1995; Ma et al., 2007). Although specified by convolving the stimulus paradigm with hemodynamic response function (HRF) for task-induced connectivity studies, the reference function is much more difficult to be constructed for resting-state fMRI studies due to the very limited prior knowledge. A common practice is then using the averaged time course of a preselected brain region of interest (ROI) as the reference function. Thus, the averaged shape, signal noise ratio (SNR), cluster's size and location of the seeds will affect the reference function and, consequently, the outcomes of the resulting connectivity map. The comparison of results using different reference vectors may present an opportunity for more detailed investigation of fMRI connectivity.

Complementary methods, driven by the data, do not make any assumption about the causes of responses or the form of the hemodynamic response function (HRF). They have been applied to functional connectivity experiments in the context of principal component analysis (PCA) (Friston et al., 1993), independent component analysis (ICA) (Ven et al., 2004), and clustering analysis (Cordes et al., 2002). These methods emphasize the intrinsic structure of the data. Among multivariate projection approaches, PCA and ICA are data-driven methods without requirement of prior information about experiment. They are based on the assumption that the activation is orthogonal or independent to the other signal variations such as brain motion, physiological signals (heart beat and respiration). Nevertheless, there are limitations to purely data-driven approaches. It is difficult to put these approaches into a statistical framework that allows one to test the activation networks against a desired hypothesis, and it lacks the ability to assess the local or regionally-specific nature of brain responses. Several groups have attempted to address the above issues. McKeown (2000) proposed a method named 'HYBICA' that allows one to use a priori hypotheses to guide the fMRI analysis. This approach combines independent components to construct task-related components and then turns to a fully hypothesis-driven approach. Hu et al. (2005) have introduced a unified method which combines temporal ICA and statistical parametric mapping (SPM) for fMRI analysis. In the present work, we propose a similar but simpler approach that combines PCA with a hypothesis-driven approach to detect functional connectivity between brain regions.

Regression analysis is a technique used for the modeling and analysis of numerical data consisting of values of a dependent variable (response variable) and of one or more independent variables (explanatory variables). It is a powerful tool for the analysis of functional connectivity mapping (Friston et al, 1995; Greicius et al., 2003). In this paper, we describe a unified method, which combines PCA, regression analysis, and statistical analysis for analyzing the functional connectivity from fMRI time-course dataset. Firstly, the principal components of the time courses of the ROI regions selected by prior information are treated as the reference functions. To obtain subject-specific connectivity maps for ROIs, regression analysis is then performed between the specific first principal component with the largest variability and time series from the whole resting-state brain on a voxel-wise basis. Finally, *t*-test is applied to weighted regression coefficient. To our best knowledge, such unified method has not been employed.

We began by presenting the theory and procedure, followed by stimulation and in vivo human fMRI experiment. We then evaluated the effectiveness of the proposed procedure. Based on the resultant accuracy, true positive rate (TPR) and receiving operator characteristic (ROC) graphs, we described how simulated fMRI data could be used to evaluate the performance of our proposed approach and compared it with the average signal based connectivity method, ICA, and PCA. Meanwhile, we described how the new approach was further tested by applying it to human data for the analysis of task-induced connectivity and the resting-state brain 'default mode' networks.

Materials and Methods

Principal Component Analysis

For fMRI, the mathematical formalism of principal component analysis is most conveniently based on the use of matrix notation to represent the data structure. Let the data set collected from the ROI be represented by the entries of a matrix X of dimension $n \times p$. The row n is the temporal dimension and the column p is the voxel-space dimension. From the theory of singular value decomposition (SVD), it is possible to find r orthogonal, $p \times 1$ principal eigenvectors $\{w_k\}$ of X^TX and r orthogonal, $n \times 1$ eigenvectors $\{u_k\}$ of XX^T, where r is the rank of the data matrix X and $r \leq \min\{n, p\}$. That is,

$$X^TXw_k = l_kw_k, \quad k = 1, 2, \ldots, r \tag{1a}$$

$$XX^Tu_k = l_ku_k, \quad k = 1, 2, \ldots, r \tag{1b}$$

where $\{l_k\}$ is the set of r non-zero eigenvalues ranked in descending order. Given the SVD of the $n \times p$ data matrix X as the spectral decomposition

$$X = ULW^T = \sum_{k=1}^{r} \sqrt{l_k}u_kw_k^T \tag{2}$$

where U is $n \times r$ with diagonal entries $l_k^{1/2}$, and W is $p \times r$. The singular values $L = diag(l_k)$ $(l_1 \geq l_2 \geq \cdots \geq l_k \geq \cdots \geq l_r \geq 0)$ are square roots of the eigenvalues $\{l_k\}$.

The orthogonal temporal principal patterns is then denoted as

$$C_{n\times p} = u_{n\times n} \times X_{n\times p} \tag{3}$$

The component with the maximal eigenvalue is considered as reference signal.

Regression Analysis of the fMRI Signal

The combined fMRI response is modeled as the linear model (Friston et al, 1995):

$$Y_i = X_i'\beta + \varepsilon_i = [X_{1i}, \cdots, X_{10i}] \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{10} \end{bmatrix} + \varepsilon_i \quad (4)$$

where β is regression coefficient, ε_i is random error, X_i is design matrix, X_{1i} is the principal component C with maximal variance, $X_{2i} = 1$, $X_{3i} = i$, $X_{4i} \cdots X_{9i}$ are 6 head motion parameters obtained in the realigning step (see below), X_{10i} is global signal obtained by averaging the time series of all the voxels within the brain.

For the first-order autoregressive (AR) model, the noise can be modeled as

$$\varepsilon_i = \rho\varepsilon_{i-1} + \zeta_{i1} \quad (5)$$

$$\widetilde{X}_1 = \sqrt{1-\rho^2}X_1, \widetilde{X}_i = X_i - \rho X_{i-1} \quad (6a)$$

$$\widetilde{Y}_1 = \sqrt{1-\rho^2}Y_1, \widetilde{Y}_i = Y_i - \rho Y_{i-1} \quad (6b)$$

where $\zeta_{i1} \sim N(0, \sigma_1^2), \rho = \frac{\sum_{i=2}^{n} \varepsilon_i \varepsilon_{i-1}}{\sum_{i=2}^{n} \varepsilon_{i-1}^2}, i = 2, \cdots, n$, so that the model becomes

$$\widetilde{Y}_i = \widetilde{X}_i'\beta + \zeta_i \quad (7)$$

where $\zeta_i \sim N(0, \sigma^2), \sigma^2 = \sigma^2/(1-\rho^2)$.

Estimating Parameters and Statistical Analysis

Denoting the pseudoinverse by $^+$, the least squares estimator of β in (7) is

$$\widehat{\beta} = \widetilde{X}^+\widetilde{Y} \quad (8)$$

where $\widetilde{X} = (\widetilde{X}_1, \cdots, \widetilde{X}_n)', \widetilde{Y} = (\widetilde{Y}_1, \cdots, \widetilde{Y}_n)'$.

Denoting the null hypothesis as

$$H_0 : A\widehat{\beta} = 0 \quad (9)$$

where $A = [1, 0, 0, 0, 0, 0, 0, 0, 0, 0]$, the residual deviation is denoted by

$$RSS = \left\| \widetilde{Y} - \widetilde{X}\beta \right\|^2, \widehat{\sigma}^2 = RSS/\upsilon \quad (10)$$

where $\upsilon = \text{rank}(\widetilde{X})$. The t-test statistic is

$$T = \frac{A\widehat{\beta}}{\|AX^+\|\widehat{\sigma}} \quad (11)$$

A random-effect (Holmes and Friston, 1998) one-tailed one-sample t-test is performed to obtain within-group positive and negative connectivity maps. The random-effect analysis estimates the error variance across subjects, rather than across cans, and thus allows the present results to be generalized to normal populations. The resulting maps of activated voxels are thresholded at a P-value, corrected for multiple comparisons using the false discovery rate (FDR)-criterion (Genovse et al. 2002). Statistical maps are displayed by superimposing the statistical results onto a standard template (MNI single T1-weighted image) as well as rendering these results onto a 3D brain reconstruction provided by SPM2 (http://www.fil.ion.ucl.ac.uk/spm).

Simulation

A simulation was conducted to evaluate the performance of our approach compared with an average signal based regression analysis approach, ICA and PCA, respectively. A slice of resting-state EPI scans (53 × 63 voxels) was replicated 200 time to simulate 200 time points of noise-free fMRI data. Nine 8 × 8 square blocs of voxels were selected for the simulation of activation (Fig. 1A). The simulated temporal signals used in this section are shown in Fig. 1B. Three simulated signals (Signals a–c) were constructed to represent the brain hemodynamics for event-related activation (Signal a), resting-state activities (Signal b) and activation in block-designed paradigm (Signal c). A slowly varying global baseline (Signal d) was added to all the voxels. To simulate the noisy environment in the brain, random noise and structured noise (Thomas et al., 2002) were mixed to the simulated data (all voxels in the brain area). The random noise (Signal e) follows a Gaussian distribution with a mean of 0 and a variance of 1. To simulate the structured noise, a cardiac signal (signal f) which has an average frequency of 1.2 Hz was generated. The magnitudes of the signals and the noises have been varied to adjust the contrast-to-noise ratio (CNR $\equiv \Delta S/\sigma_{noise}$). The CNRs ranged from 0. 75 to 3, consistent with CNR values reported in the fMRI literature (Esposito et al. 2002; Huettel et al. 2004).

Participants

Forty eight healthy right-handed volunteers [21 females, age range: 17–36 years, mean (SD): 25.3 (4.76)] with no neurological and psychiatric disorder participated in this study. All MR examinations were carried out under the guidance of the Declaration of Helsinki 1964. Written informed consents were also obtained from all participants. This study was approved by the local Medical Ethics Committee at Jinling Hospital.

Experimental Paradigms

Two fMRI experiments were conducted. In the first experiment, the subjects were scanned during a resting state. All the healthy subjects participated in the first experiment and they were instructed to relax but remain awake, keep their eyes closed and restrain from any cognitive, language, or imaginary visual or motor task as much as possible during the functional scan. The purpose of this

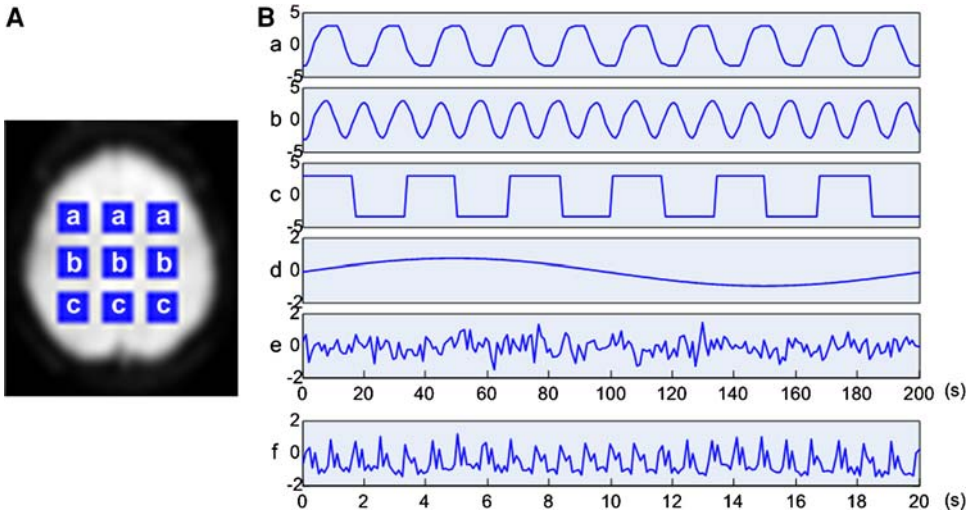

Fig. 1 **A** An EPI scan image with the selected nine regions of interest and simulated fMRI time courses **B**. Signals a and b (0.08 Hz) are the series of gamma variate functions simulating event-related and resting-state brain hemodynamics, respectively. Signal c is the convolution of an HRF and a square wave simulating a block-design fMRI signal. Signal d (0.005 Hz) is sine wave simulating a slowly varying global baseline. Signal e is a Gaussian signal simulating the random white noise. Signal f (with a mean frequency of 1.2 Hz) is a cardiac signal simulating the structured noise

experiment is to verify that the new proposed method can be used to map functional connectivity in the regions of the default mode, the well defined functional brain networks during resting-state (Raichle 2001; Greicius et al. 2003; Mantini et al. 2007). In the second experiment, 12 of the healthy subjects were scanned when performing a hand flexion task using their non-dominant (left) hand (Wu et al, 2008). The subjects were trained to grip the hand with a frequency of 1 HZ and practiced 100 s before the scan. A block design was used in the paradigm and the overall task consisted of 20 blocks, 10 task blocks alternating with 10 resting blocks lasting for 400 s with each block of 20 s. During the functional scan, the subjects were instructed to grip the left hand when seeing a stationary cross presenting on the center of the screen throughout each task block, and to remain still and fixate on a stationary asterisk throughout each resting block. The purpose of this experiment is to obtain a seed region from actual activations induced by a hand movement task, and further compare the proposed method with the classical method when disentangling brain motor networks.

MRI Data Acquisition and Preprocessing

Imaging data were acquired using a 1.5-Tesla scanner (GE-Signa, Milwaukee, US.). The fMRI time series of 200 repeated whole brain images were acquired in an orientation parallel to the AC-PC plane using a T_2^*-weighted GRE-EPI sequence. The sequence parameters were: TR = 2000 ms, TE = 40 ms, FA = 80°, FOV = 24 × 24 cm^2, 23 continuous slices with a thickness of 4 mm (no gap), matrix = 512 × 512. Anatomical images were acquired in the same orientation with a T1-Flair sequence (TR = 2019.3 ms, TE = 25.3 ms, interslice gap = 0.5 mm, slice thickness = 4 mm, matrix = 64 × 64). Each subject laid inside the scanner with foam pads so as to prevent from head movement.

Data preprocessing was partly performed using the software package of SPM2. After slice timing, realignment for head motion correction, four subjects whose head motion exceeded ±0.5 mm and ±0.5°during scanning, were excluded. The standard Montreal Neurological Institute (MNI) template provided by SPM was used in normalization with resampling voxel size of 3 × 3 × 3 mm^3. After being smoothed with FWHM (full width at half maximum) of 8 mm kernel, the data were temporally filtered (band pass, 0.01 ~ 0.08 Hz) to remove the effects of very-low-frequency drift and high frequency noise by using finite impulse-response (FIR) filter (Biswal et al. 1995). To perform the functional connectivity analysis, time series from the resting-state scan were extracted for the subject-specific ROIs. Before extracting the references, scaling step was performed across the ROI voxels. To minimize the effect of global drift, each column of data matrix X was normalized to a mean of zero (Friston et al. 1994, 1995).

Selection of ROIs

For the simulated data, the simulated signals in every three ROIs (Fig. 1Aa, b, and c) were identical. Every three ROIs were chosen as a volume of seeds for both the PCA based and average signal based connectivity methods. For the

human resting-state data, two different ROIs (ROI I and II) were selected for the proposed method. ROI I contains the voxels inside a spherical ROI (radius = 10 mm) in the posterior cingulated cortex (PCC)/precuneus area (0, −56, 30) (Fransson 2005), which were extracted from the resting-state scan in all subjects. Previous work has shown that seeding the PCC/precuneus produces the most comprehensive map of the default mode network (Greicius et al. 2004). ROI II was chosen based on the motor cortex activations detected using the data from the hand movement task. For the task-induced human data, *t*-test was used to determine the main effect of movement, the activated areas in the whole brain were highlighted ($P = 0.001$, uncorrcted). The ROI II was then defined as a sphere within the activated area located in the contralateral primary motor cortex (M1) corresponding to the dominant hand (right hand), centered at the voxel of the highest statistical value and with a radius of 10 mm. The principle component of the time courses of the seed regions were treated as the reference function for our simulation and resting-state connectivity network analysis.

Results

Simulation Results

The performances of PCA based connectivity method and other conventional functional connectivity methods such as average signal based connectivity, ICA and PCA methods, were first investigated using the simulated data. To test reliability, we repeated the entire procedure 20 times. The extractions of spatial components by means of ICA and PCA were described in detail in the literatures (Baumgartner et al. 2000; Bell et al. 1995; Friston et al. 1993; Ma et al. 2007). Both ICA and PCA were applied directly to the simulation data. All the resulting values calculated by the four connectivity methods were converted to their corresponding *Z* scores (Ma et al. 2007; Zhao et al. 1999). Significant differences between the four performances are shown at a threshold ($P < 0.01$, corrected). The effects of noise levels on the detection of functional connectivity networks using these methods are shown in Figs. 2(a–d), respectively. In each case, four CNRs (CNR = 0.75, CNR = 1, CNR = 2 and CNR = 3) were evaluated.

It is noticeable that the PCA only method gives the worst description of the dataset in terms of extracting each spatial network (see Fig. 2d). The main three components (Fig. 2m–p) extracted in different CNRs yield two source signals (e.g., Fig. 1a and b). This indicates that the PCA is an effective dimension reduction method for classification, but performs poorly in unmixing the spatial components. The comparisons of the spatial maps provided by the remaining methods show that all the three methods perform well in detecting the signals when the noise level was low (CNR = 2, 3). When the noise was high (CNR = 0.75, 1), PCA based approach had a better performance for detecting the signal. For the two hypothesis-driven method, the PCA based approach was hardly affected by the noise (Fig. 2a and b), whereas average signal based approach was significantly affected by the noise (Fig. 2e and f.). In each CNR, our proposed method can detect the signals with a highest true positive rate (TPR) (Fig. 3).

The accuracy and receiving operator characteristic (ROC) curves were used for further comparison using simulated data for different *P* levels. The results shown in Fig. 4 indicate that under almost all the conditions, the PCA based connectivity approach has higher accuracy than the other three connectivity methods, which is evident especially when the CNRs is high (CNR = 2, 3). The ROC

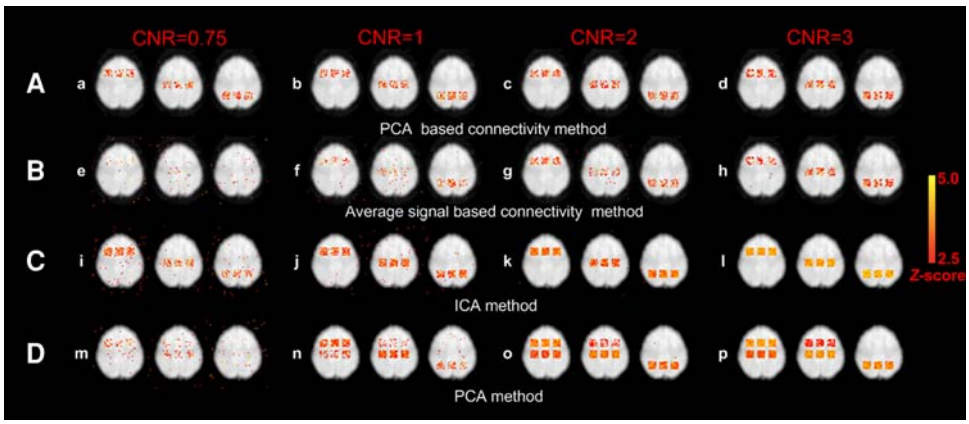

Fig. 2 Functional connectivity networks detected by PCA based method, average signal based method, ICA, and PCA under different CNRs. **A** PCA based connectivity results. **B** Average signal based connectivity method. **C** Purely data-driven ICA method. **D** Purely data-driven PCA method. Note that the purely data-driven PCA method gives the worst description of the results in terms of extracting each spatial network independently

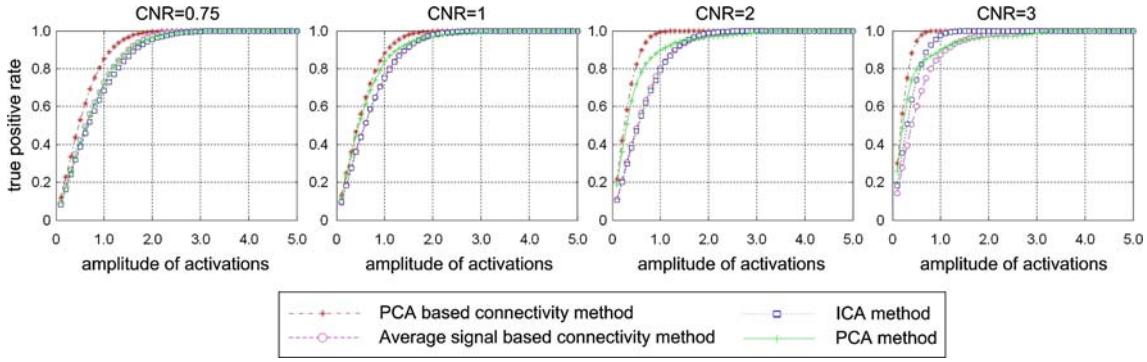

Fig. 3 The true positive rate (TPR) of PCA based connectivity method, average signal based connectivity method, ICA and PCA at CNRs of 0.5, 1, 2 and 3. The plots are the average TPR curves of the 20 repeated procedures

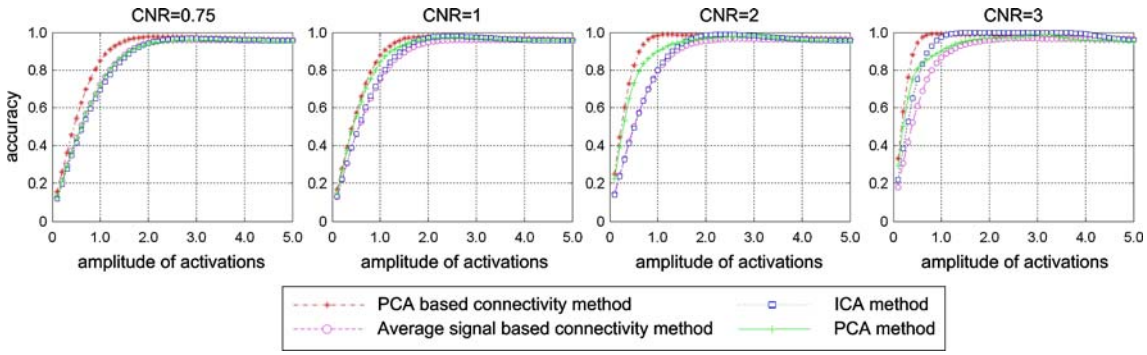

Fig. 4 Accuracy analyses of PCA based connectivity method, average signal based connectivity method, ICA and PCA at CNRs of 0.75, 1, 2 and 3. The plots are the average accuracy curves of the twenty repeated procedures

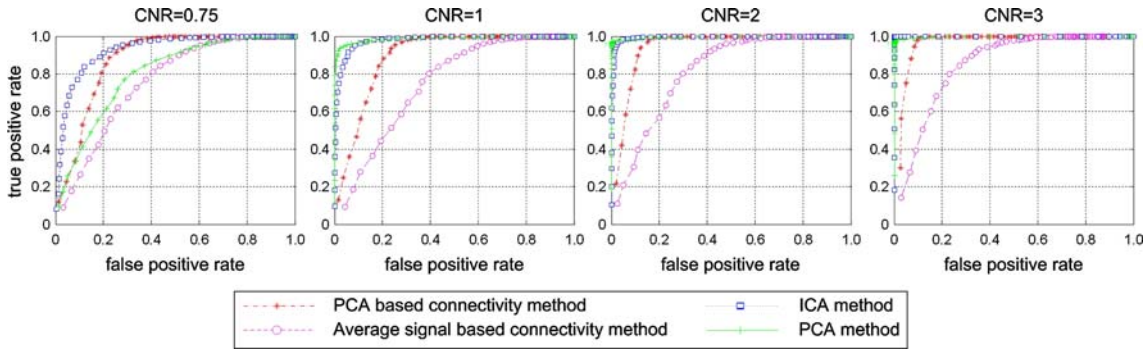

Fig. 5 The ROC curves for PCA based connectivity method, average signal based connectivity method, ICA and PCA at CNRs of 0.75, 1, 2 and 3. Plotted are the average ROC curves of 20 repeated procedures

curves shown in Fig. 5, which indicated the true-positives vs. the false-positives rates, also indicate that the PCA based connectivity approach outperforms the average signal based approach. However, when CNRs is high (CNR = 1, 2, 3), the data-driven methods are better than our proposed method and average signal based connectivity method in detecting the activity as described in the literature (Ma et al. 2007). The area under the ROC curve (AUC) was taken as a scalar measure (Hanley and McNeil 1982). An area of 1 represents a perfect test. For the two hypothesis-driven methods, the AUCs of PCA based connectivity method are 0.8828, 0.8980, 0.9473 and 0.9713 respectively, which are higher than the AUCs (0.7604, 0.7795, 0.8181 and 0.8528) of average signal based approach (Fig. 6).

Identification of Resting-State Networks

The pattern of resting-state brain network in the healthy subjects was identified in order to further validate the new approach. Based on the ROI I, the functional connectivity

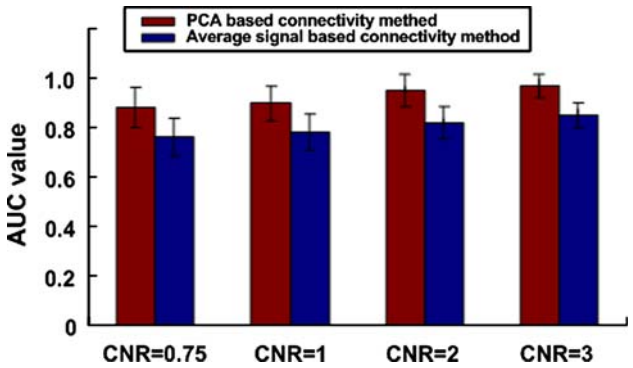

Fig. 6 AUC values of ROC curves (mean ± SD, $n = 20$). For the simulation with a CNR set to be 0.75, the AUC is 0.8828 and 0.7604 for PCA based method and average signal based method, respectively. Briefly, CNR = 1, AUC = 0.8980 (PCA based) and 0.7795 (average signal based); CNR = 2, AUC = 0.9473 (PCA based) and 0.8181 (average signal based); CNR = 3, AUC = 0.9713 (PCA based) and 0.8528 (average signal based)

maps obtained by the proposed approach in 48 healthy subjects during rest are shown in Fig. 7. A 3D brain reconstruction of the results is also shown in Fig. 8. A seed reference time course was obtained by PCA courses within the ROI. Regression analysis was carried out between the seed reference and the whole brain in a voxel-wise manner. The *t*-test was applied to improve the normality of these regression coefficients. The individual t-value was then entered into a random effect one-sample *t*-test to determine brain regions showing significant connectivity to the ROI within each group. The threshold was set at $P < 0.01$, corrected.

During rest and independent of resting-state condition, a full description of all regions that correlated positively and negatively with the PCC/precuneus is given in Table 1. Consistency was fairly good across the subjects. The two networks of spontaneous low-frequency fluctuations from the multi-subject PCA based regression analysis have been cited in numerous studies to increase during resting state and suspend during specific goal-directed behaviors (Fransson 2005; Garrity et al. 2007; Raichle et al. 2001). This confirms previously published results using seed-based functional connectivity measures, and provides further evidence that the PCC/precuneus connectivity is involved in the neuropathology of resting-state human brain.

The motor networks detected by the PCA based and average signal based methods were also evaluated using the resting-state fMRI data based on the ROI II. Group results are shown in Fig. 9, respectively. For the motor system, the detected regions were M1, primary somatosensory cortex (S1), and the supplementary motor area (SMA). The maps produced with the two processing methods were similar in most cases. However, the spatial regions mapped by PCA based method were more specific than those by the average signal based connectivity method.

Discussion

We have introduced an approach for detecting the functional connectivity based on PCA, regression analysis and *t*-test analysis. The proposed algorithm has a number of

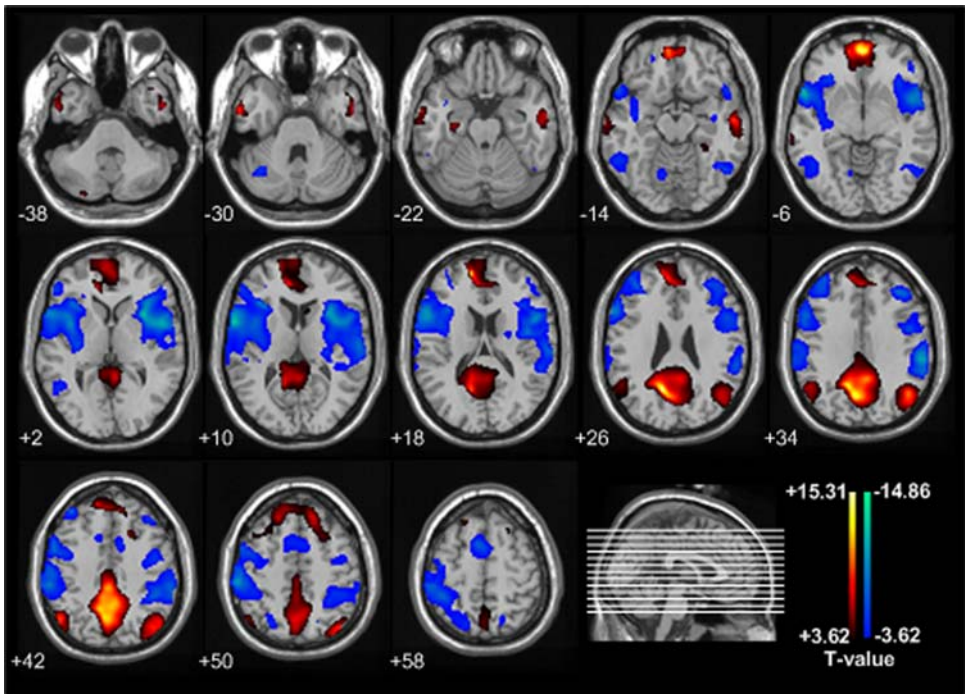

Fig. 7 Functional connectivity maps calculated by the PCA based regression analysis approach in 48 healthy subjects during rest. Brain regions color-coded in a red–yellow color scale correlate positively with the PCC/precuneus region and the areas color-coded in blue-magenta color scale correlate negatively with the ROI region. Two statistical *T*-maps ($P < 0.01$, FDR corrected) superimposed on a single T1-weighted subject image available in SPM2 are shown. The numbers beneath the axial MR image refer to MNI coordinates

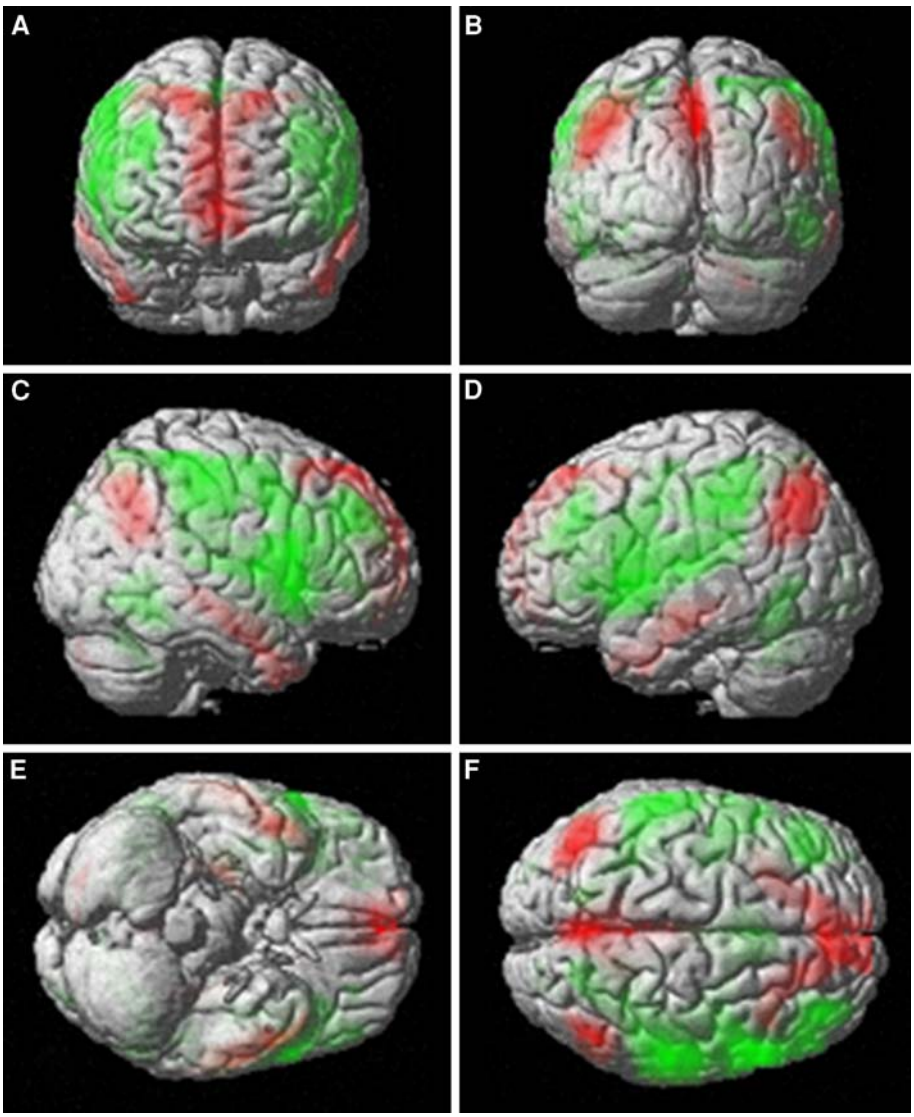

Fig. 8 Brain regions with significant positive connectivity (*red*) and negative connectivity (*green*) ($P < 0.01$,FDR corrected) in 48 healthy subjects during rest are shown on a surface rendering of an MRI of the brain. **a** anterior view, **b** posterior view, **c** contralateral view, **d** ipsilateral view, **e** inferior view, **f** superior view

advantages over the average signal based approach. Simulation results illustrate three key points. First, in the low-noise case (Fig. 2c, d, g and h, CNR = 2 and 3), both the approaches do well, with the PCA based approach slightly outperforming average signal based approach. Secondly, in the high-noise case (Fig. 2 a, b, e and f, CNR = 0.5 and 1), the PCA based approach was found to be more robust whereas the average signal based approach show considerable dispersion. Thirdly, in both the noise situation, the PCA based approach shows improved accuracy and AUC value over average signal based method (Figs. 4 and 6). These results suggest that the proposed approach outperforms the traditional voxel-based method and is more appropriate for the analysis of brain functional connectivity analysis.

Another aim of the study therefore was to carry out a functional connectivity analysis of the networks of regions involved in sustaining a resting state and a motor system in the brain. Taking into account the effect of seeds on connectivity maps, we attempted to avoid biased choice of networks by selecting two different ROIs in the analyses. For the ROI I, the detected connectivity in the resting-state fMRI datasets by the proposed method contained most of the significant activation which was described as a default mode (Greicius et al. 2003; Raichle 2001) and resting-state networks (RSNs) (De Luca et al. 2006; Mantini et al. 2007) of brain function. Also, the activation maps obtained with the approach contained less spurious activation and noise. A comprehensive and qualitative comparison of the 'default mode' by means of the PCA based method and other conventional connectivity methods is impossible because the gold standards of the connectivity activation patterns for individual subjects are unavailable. In addition, we have provided the ROI II from a task-related fMRI experiment for evaluating the proposed method. When the reference is extracted through principle component, the

Table 1 Functional connectivity detected by PCA based method with the precuneus/PCC (ROI) during the resting state

Regions	MNI x, y, z	BA	Peak *t*-value
Positive PCC/precuneus connectivity			
Precuneus (R)	12 −60 27	31	15.31
PCC (L)	−6 −48 39	31	13.73
MPFC (L/R)	−2 54 −6/12 54 21	10/9	14.26/12.66
vACC	3 48 0	32	4.99
Middle temporal gyrus (L/R)	−57 −12 −12/63 −15 −18	21	11.87/6.18
Inferior temporal gyrus (L/R)	−51 0 −39/51 0 −39	20	7.66/7.99
Angular gyrus (L/R)	−45 −63 36/51 −57 36	39	11.15/8.43
Temporal fusiform gyrus(R)	57 −3 −30	20	9.40
Parahippocampal gyrus (R)	27 −18 −24	–	6.97
Cerebellum posterior lobe (R)	27 −84 −39	–	6.07
Caudate (L)	−9 15 9	–	5.42
Negative PCC/precuneus connectivity			
Ventral PM (L/R)	−57 6 15/60 6 12	44	−11.44/−13.01
SMA	6 9 57	6	−8.52
DLPFC (L/R)	−36 36 30/36 39 30	10	−9.97/−8.63
Dorsal PM (L/R)	−51 3 30/57 6 33	9	−8.02/−7.76
IPL (L/R)	−57 27 33/60 −30 36	40	−11.19/−7.02
Postcentral gyrus (L/R)	−57 −18 36/57 −27 48	4/2	−7.92/−10.74
Middle temporal gyrus (L/R)	−57 66 −6/51 −51 −12	37	−5.21/−5.86
Middle occipital gyrus (L/R)	−54 −60 −9/51 −60 −9	19	−4.87/−5.02
Lingual gyrus (L/R)	−15 −75 −12/15 −70 −9	18	−3.78/−3.71
Insula (L/R)	−33 12 6/42 12 0	13	−14.86/−8.95
Cerebellum posterior lobe (L/R)	30 66 −30/30 −63 −30	–	−3.62/−3.83

PCC posterior cingulated cortex, *ROI* region of interest, *BA* Brodmann area, *L* left, *R* right, *PM* premotor cortex, *SMA* supplementary motor area, *MPFC* medial prefrontal cortex, *DLPFC* dorsolateral prefrontal cortex, *vACC* ventral anterior cingulate cortex, *IPL* inferior parietal lobule

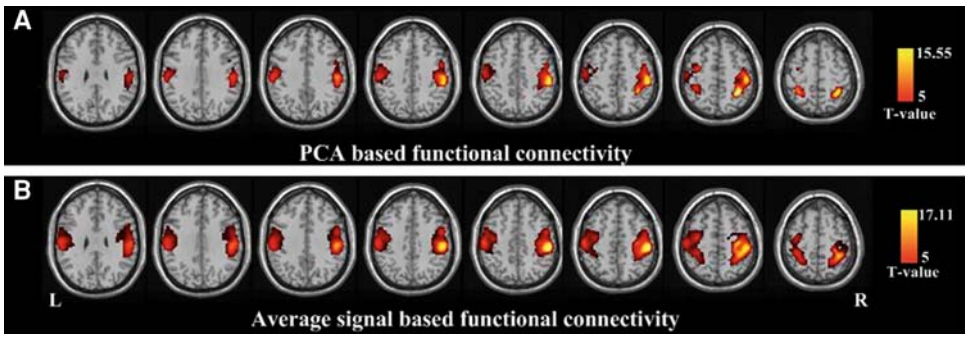

Fig. 9 The motor functional networks calculated by the PCA based regression analysis approach and the average signal based regression analysis approach in 12 healthy subjects during rest. **a** motor networks detected by PCA based method. **b** motor networks detected by average signal based method

calculated motor networks from the statistical results become more specific than those using average signal reference.

Traditionally, averaging across repeated observations has been the common method of choice for improving signal-to-noise and in turn the detectability of an effect in fMRI activation or connectivity studies. Likewise, the joint implication of a collection of voxels with the same time-course behavior improves greatly on the statistical significance as compared to the analysis of the response for a single voxel out of context. Whereas, the technique is potentially problematic because the hemodynamic response may varies across the brain. Such averaging may destroy significant features in the temporal evolution of the fMRI response that stem from either difference in vascular coupling to neural tissue or actual differences in the neural response between the averaged voxels. Averaging signals are strongly affected by the seeds location, cluster's size and may only detect partial systems (Ma et al. 2007). It is seeds sensitive. In comparison, PCA is a data-driven method for seeking out intrinsic structure, each implying an eigenimage associated with a unique temporal response profile. In comparison to univariate average signal based methods where the reference vector represents an a priori fixed effect or the temporal response within a manually selected region of interest, the reference or template time

course vector in PCA is a unique feature of the data itself. PCA also acts as a variation reduction technique, relegating most of the random noise to the trailing components while collecting systematic structure into the leading ones. On the whole, the functional connectivity processing of PCA based approach and average signal based method are similar but not identical. The principal eigenvectors capture both the spatial and temporal aspects of fMRI data, and the first temporal principal eigenvector is inherently matched to unique and uncorrelated features and are ranked in order of the amount of variance explained.

The purely data-driven methods ICA, PCA and our proposed PCA based connectivity approach are both capable of detecting functional connectivity networks (Johnson and Wichern 1998; Friston et al. 1993; McKeown et al. 1998a). However, the performances of them are not necessarily the same. They may significantly differ from the results in terms of the size and location of the detected networks. The ICA and PCA methods were applied to the same simulated data set and the results were compared (Figs. 2, 3, 4 and 5). Based on our results, the overall performance of the proposed method appears to be better than that of ICA and PCA in terms of accuracy and TPR. However, the resulting ROC curves indicate that the data-driven method have a better performance in terms of specificity. The limitation of the purely data-driven method is difficult to associate thresholds with significant levels. As a technique of blind source separation, PCA separates signals by utilizing the orthogonality between each component, and ICA is used in signal processing contexts to separate out mixtures of independent sources. Since the amplitude of a separated component is arbitrary, it is theoretically difficult to assess the statistically significant level of the map (McKeown et al. 1998b; Sychra et al. 1994). The z-map conversion of data-driven methods tend to overestimate the false-positive rate (FPR) (Ma et al. 2007; McKeown et al. 1998a).Our proposed approach which combines the PCA, regression analysis and *t*-test, is significantly easy to put these approaches into a statistical framework that allows one to test the activation networks against a desired hypothesis. The resulting maps of activated voxels in the resting-state connectivity networks can be thresholded and corrected for multiple comparison using the FDR-criterion. Because our proposed method is based upon an explicit statistical model, it does not suffer from the disadvantages of ICA and PCA.

Conclusion

It is widely recognized that clarifying neural connectivity is essential for the understanding of brain function. In this paper, we have advocated a new approach combining PCA, regression analysis and statistic analysis to fMRI signal analysis in the study of brain functional connectivity. Based on our simulated fMRI data, the curves of accuracy and TPR indicated that the proposed method outperforms those previously described connectivity methods. Meanwhile, such combined use of the local PCA and GLM indeed has not been reported in the literature. Based on in vivo fMRI data, results show similar associations consistent with the published results using seed-based functional connectivity measures implicated in resting-state networks of the brain 'default mode'. Compared with the average signal based approach, the PCA based approach provides a flexible way to analyze the connectivity of fMRI data that increases the SNRs of ROI seed and obtains more specific motor networks. The use of PCA in a connectivity analysis, which combines elements of both regression approach and statistic analysis approach, may be proved to be a useful tool for fMRI connectivity analysis.

Acknowledgment This work was supported by the Nature Science Foundation of China (30470510, 30670600, and 30800264; 60628101).

Refenrences

Baumgartner R, Ryner L, Richter W, Summers R, Jarmasz M, Somorjai R (2000) Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering vs. principal component analysis. Magn Reson Imaging 18:89–94

Bell AJ, Sejnowski TJ (1995) An information-maximization approach to blind separation and blind deconvolution. Neural Comput 7:1004–1034

Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Res Med 34:537–541

Cordes D, Haughton V, Carew JD, Arfanakis K, Maravilla K (2002) Hierarchical clustering to measure connectivity in fMRI resting-state data. Magn Reson Imaging 20(4):305–317

De Luca M, Beckmann CF, De Stefano N, Matthews PM, Smith SM (2006) fMRI resting state networks define distinct modes of long-distance interactions in the human brain. Neuroimage 29(4):1359–1367

Esposito F, Formisano E, Seifritz E, Goebel R, Morrone R, Tedeschi G, Di Salle F (2002) Spatial independent component analysis of functional MRI time-series: to what extent do results depend on the algorithm used? Hum Brain Mapp 16:146–157

Fransson P (2005) Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of resting-state default mode of brain function hypothesis. Hum Brain Mapp 26:15–29

Friston KJ (1994) Functional and effective connectivity in neuroimaging: a synthesis. Hum Brain Mapp 2:56–78

Friston KJ, Frith CD, Liddle PF, Frackowiak RS (1993) Functional connectivity: the principal-component analysis of large (PET) data sets. J Cereb Blood Flow Metab 13(1):5–14

Friston KJ, Holmes AP, Poline JB, Grasby BJ, Williams CR, Frackowiak RSJ, Turner R (1995) Analysis of fMRI time-series revisited. Neuroimage 2:45–53

Garrity AG, Pearlson GD, McKiernan K, Lloyd D, Kiehl KA, Calhoun VD (2007) Aberrant "default mode" functional connectivity in schizophrenia. Am J Psychiatry 164:450–457

Genovse CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging: using the false discovery rate. Neuroimages 15:870–878

Greicius MD, Krasnow B, Reiss AL, Menon V (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci USA 100(1):253–258

Greicius MD, Srivastava G, Reiss AL, Menon V (2004) Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc Natl Acad Sci USA 101(13):4637–4642

Hanley J, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1):29–36

Holmes AP, Friston KJ (1998) Generalisability, random effects and population inference. Neuroimage 7:S754

Hu D, Yan L, Liu Y, Zhou Z, Friston KJ, Tan C, Wu D (2005) Unified SPM-ICA for fMRI analysis. Neuroimage 25:746–755

Huettel SA, Song AW, McCarthy G (2004) Functional magnetic resonance imaging Sinauer Associates: Sunderland, p xviii, 492

Johnson RA, Wichern DW (1998) Applied multivariate statistical analysis, 4th edn edn. Prentice-Hall, Inc., Upper Saddle, NJ

Lowe MJ, Mock BJ, Sorenson JA (1998) Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations. Neuroimage 7:119–132

Ma L, Wang B, Chen X, Xiong J (2007) Detecting functional connectivity in the resting brain: a comparison between ICA and CCA. Magn Reson Imaging 25(1):47–56

Mantini D, Perrucci MG, Del Gratta C, Romani GL, Corbetta M (2007) Electrophysiological signatures of resting state networks in the human brain. Proc Natl Acad Sci USA 104:13170–13175

McKeown MJ (2000) Detection of consistently task-related activations in fMRI data with hybrid independent component analysis. Neuroimage 11:24–35

Mckeown MJ, Sejnowski TJ (1998) Independent component analysis of fMRI data: examining the assumptions. Hum Brain Mapp 6:368–372

Mckeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ (1998) Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp 6:160–188

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. Proc Natl Acad Sci USA 98:676–682

Strother SC, Anderson JR, Schaper KA, Sidtis JJ, Liow JS, Woods RP, Rottenberg DA (1995) Principal component analysis and the scaled subprofile model compared to intersubject averaging and statistical parametric mapping: I. ''Functional connectivity'' of the human motor system studied with [^{15}O] water PET. J Cereb Blood Flow Metab 15:738–753

Sychra JJ, Bandettini PA, Bhattacharya N, Lin Q (1994) Synthetic images by subspace transforms I. Principal components images and related filters. Med Phys 21(2):193–201

Thomas CG, Harshman RA, Menon RS (2002) Noise reduction in BOLD-based fMRI using component analysis. Neuroimage 17(3):1521–1537

Ven VG, Formisano E, Prvulovic D, Roeder CH, Linden DEJ (2004) Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. Hum Brain Mapp 22(3):165–178

Wu X, Chen K, Liu Y, Long Z, Wen X, Jin Z, Yao L (2008) Ipsilateral brain deactivation specific to the nondominant hand during simple finger movements. Neuroreport 19(4):483–486

Zhao X, Glahn D, Tan L, Li N, Xiong J, Gao J (1999) Comparison of TCA and ICA techniques in fMRI data processing. J Magn Reson Imaging 19:397–402