RV-coefficient and its significance test in mapping brain functional connectivity

Article in Proceedings of SPIE - The International Society for Optical Engineering · February 2009
DOI: 10.1117/12.811369

CITATION
1

READS
49

4 authors:

Hui Zhang
Chinese Center For Disease Control And Preven...
921 PUBLICATIONS 30,493 CITATIONS
SEE PROFILE

Jie Tian
Chinese Academy of Sciences
661 PUBLICATIONS 7,114 CITATIONS
SEE PROFILE

Jun Li
Xidian University
20 PUBLICATIONS 220 CITATIONS
SEE PROFILE

Jizheng Zhao
Northwest A & F University
6 PUBLICATIONS 88 CITATIONS
SEE PROFILE

All content following this page was uploaded by Jie Tian on 12 October 2015.
The user has requested enhancement of the downloaded file. All in-text references underlined in blue are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.
RV-coefficient and its significance test in mapping brain functional connectivity

Hui Zhang\textsuperscript{a}, Jie Tian\textsuperscript{*a,b}, Jun Li\textsuperscript{b} and Jizheng Zhao\textsuperscript{b}

\textsuperscript{a}Medical Image Processing Group, Key Laboratory of Complex Systems and Intelligence Science, Institute of Automation Chinese Academy of Science. P.O.Box 2728, Beijing, 100190, China.
\textsuperscript{b}Life Science Research Center, Xidian University, Xi’an, Shaanxi 710071, China

\*Corresponding author: Email: tian@ieee.org, jie.tian@ia.ac.cn, Telephone: 86-10-82618465; Fax: 8610-62527995

ABSTRACT

The statistic of RV-coefficient is a good substitute for the Pearson correlation coefficient to measure the temporal similarity of two local brain regions. However, the hypothesis test of RV-coefficient is a hard problem which limits its application. This paper discussed the problem in details. Since the distribution of RV-coefficient is unknown, we do not know a critical p-value to statistically test its significance. We proposed a new strategy to test the significance of RV calculated from fMRI. In order to approximate the p-value, we elicited the first two moments of the population permutation distribution of RV; we then derived a formula to more closely approximate the normal distribution with these transformed statistics. These transformations of statistics are suggested for a precise approximation to the permutational p-value even under large number of observations. This strategy of test can greatly reduce the computational complexity and avoid “calculation catastrophe”, we then use the statistic of RV to extract the map of functional connectivity from fMRI and test its significance with the strategy proposed here.

Keywords: RV-coefficient; permutation test; BOLD-fMRI; functional connectivity

1. INTRODUCTION

Blood-oxygen-level dependent functional magnetic resonance imaging (BOLD-fMRI) provides us a useful tool for the research into human brain function due to its high spatiotemporal resolution. Besides the large amount of studies on the functional region’s neural response to an external stimulus, recently, great attention has been focused on the functional connectivity of the spatially discrete regions\textsuperscript{1-11}, especially the functional connectivity of those brain regions in resting states\textsuperscript{1-9}. The traditional method for measuring the brain functional connectivity is Pearson correlation analysis\textsuperscript{1-3}. By calculating the Pearson correlation coefficient of the time course of each brain voxel with the averaged time courses in the brain region of interests (ROI), the whole brain functional connectivity map is obtained and the strength of the connectivity of local brain regions with that of the regions of interest is measured. The traditional Pearson correlation based functional connectivity is easy to calculated, but the fine-scale structure based spatial connectivity information\textsuperscript{12,13} has not been extracted due to this voxel-by-voxel univariate analysis method. Meanwhile, in order to increase the signal to noise ratio (SNR) and statistical power, the fMRI data usually spatially smoothed beforehand with a Gaussian kernel, this process greatly decrease the native resolution of the data, and the fine-scale patterns of weak connectivity that characterize neuroscientifically relevant information are blurred or discarded\textsuperscript{12}.

The statistic of RV-coefficient\textsuperscript{14-20} is a good substitute for the Pearson correlation coefficient to measure the temporal
similarity of two local brain regions; this multivariate measure has many advantages in analyzing functional connectivity from fMRI data: it is based on the regional homogeneity (ReHo) to directly measure two local brain regions; it can efficiently eliminate spatial noise and map the functional connectivity with fine-scaled pattern; However, the hypothesis test of RV-coefficient is a hard problem which limits its application\textsuperscript{15-17}. This paper discussed the problem in detail. Since the distribution of RV-coefficient is unknown, we do not know a critical p-value to statistically test its significance. Some researchers seek to permutation test to estimate the p-value, but this method does not seem applicable in fMRI due to large number of observations induced “calculation catastrophe”. Here we proposed a new strategy to test the significance of RV calculated from fMRI. In order to approximate the p-value, we elicited the first two moments of the population permutation distribution of RV; we then derived a formula to more closely approximate the normal distribution with these transformed statistics. These transformations of statistics are suggested for a precise approximation to the permutational p-value even under large number of observations. This strategy of test can greatly reduce the computational complexity and avoid “calculation catastrophe”, we then use the statistic of RV coefficient to map the functional connectivity from fMRI in resting states by using each subject’s fusiform face area as regions of interest\textsuperscript{21-24} (ROI) and test its significance with the strategy proposed here.

2. METHODS

2.1 RV-coefficient and the measure of brain functional connectivity

The statistic of RV coefficient was firstly introduced in 1976 by Robert and Escoufier\textsuperscript{14}. For measuring the temporal similarities of two local brain regions, the statistic of RV-coefficient can be denoted as:

\[
RV(X, Y) = \frac{\text{trace}(XX'^TY')}{\sqrt{\text{trace}(XX'^T) \times \text{trace}(Y^TY')}}^2
\]

(1)

Here, \(X\) and \(Y\) are \(n \times p\) and \(n \times q\) matrix from two local brain regions of fMRI data, \(X'\) is the transpose of matrix \(X\), \(Y'\) is the transpose of matrix \(Y\), \(\text{trace}(X)\) is the trace of the matrix \(X\). \(p\) and \(q\) are the number of voxels from local brain regions, \(n\) is the number of observed fMRI time point. This multivariate statistic provides a direct and precise way to extract the functional connectivity of brain network from fMRI data.

2.2 The first three moment of the permutated RV-coefficient

To test the significance of RV with large number of observed time points, a practical method is to approximate the permutation distribution of RV by a continuous distribution under the null hypothesis that data sets \(X\) and \(Y\) are independent\textsuperscript{15-17}, so we deduce the first two moments of permutation distribution of RV:

Let:

\[
X = (x_1, x_2, ..., x_n)', \ Y = (y_1, y_2, ..., y_n)'
\]

\[
A = XX'^T, \ B = YY'^T
\]
$A = XX^t = (x_1, x_2, \ldots, x_n) (x_1, x_2, \ldots, x_n) = (\ldots, x_i^t x_j, \ldots)$,

$B = YY^r = (y_1, y_2, \ldots, y_n) (y_1, y_2, \ldots, y_n) = (\ldots, y_i^t y_j, \ldots) \quad i, j \in 1, 2, \ldots n$

$X_{(X)}$ and $Y_{(Y)}$ must be mean-centered by column. Permute all rows of $Y$ to get all the value of permutated $RV_{(X,Y)}$, we denote $E_{\text{perm}}(RV_{(X,Y)})$ and $V_{\text{perm}}(RV_{(X,Y)})$ the mathematical expectation and variance of permutated $RV_{(X,Y)}$ respectively.

$$E_{\text{perm}}(RV_{(X,Y)}) = E_{\text{perm}} \left( \frac{\text{tr}(AB)}{\text{tr}(AA)^{1/2} \times \text{tr}(BB)^{1/2}} \right) = \frac{1}{\text{tr}(AA)^{1/2} \times \text{tr}(BB)^{1/2}} E_{\text{perm}}(\text{tr}(AB))$$

Let $\Omega = \{ \sigma | \sigma(i) = j \quad i, j \in \{1, 2, \ldots n\} \}$, $\sigma(i)$ is the index to perform each permutation of series $(1, 2, \ldots n)$:

$$E_{\text{perm}}(\text{tr}(AB)) = E_{\text{perm}}(\text{tr}(x_i x_j y_i y_j)) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^t x_j y_i y_j$$

The Equation 3 can be decomposed into $E_{\text{perm}} = E_{\text{perm}1} + E_{\text{perm}2}$ according to $(i = j) \cup (i \neq j)$:

$$E_{\text{perm}1} = \frac{1}{n!} \sum_{i=1}^{n} x_i^t x_i y_{\sigma(i)}^t y_{\sigma(i)} = \frac{(n-1)!}{n!} \sum_{i=1}^{n} x_i^t x_i y_j^t y_j = \frac{1}{n} \text{tr}(XX^t) \text{tr}(YY^t) = \frac{\text{tr}(A)\text{tr}(B)}{n}$$

$$E_{\text{perm}2} = \frac{1}{n!} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^t x_j \sum_{\sigma \in \Omega} y_{\sigma(i)}^t y_{\sigma(j)} = \frac{(n-2)!}{n!} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^t x_j \sum_{i \neq j} y_i^t y_j$$

Because matrix $Y$ are mean-centered by column, $\sum_{i=1}^{n} \sum_{j=1}^{n} y_i^t y_j = 0$, we have

$$E_{\text{perm}2} = -\frac{1}{n(n-1)} \text{tr}(YY^t) \sum_{i=1}^{n} \sum_{j \neq i} x_i^t x_j = \frac{1}{n(n-1)} \text{tr}(YY^t) \text{tr}(XX^t) = \frac{\text{tr}(A)\text{tr}(B)}{n(n-1)}$$

Combining the above two formula we get:

$$E_{\text{perm}}(\text{tr}(AB)) = E_{\text{perm}} = E_{\text{perm}1} + E_{\text{perm}2} = \frac{\text{tr}(A)\text{tr}(B)}{n-1}$$

$$E_{\text{perm}}(RV_{(X,Y)}) = \frac{1}{\text{tr}(AA)^{1/2} \times \text{tr}(BB)^{1/2}} = \frac{\text{tr}(A)\text{tr}(B)}{(n-1) \times \text{tr}(AA)^{1/2} \times \text{tr}(BB)^{1/2}}$$

For variance of permutated $RV_{(X,Y)}$:

$$V_{\text{perm}}(RV_{(X,Y)}) = E_{\text{perm}}(RV_{(X,Y)})^2 - (E_{\text{perm}}(RV_{(X,Y)}))^2$$

$$\left( E_{\text{perm}}(RV_{(X,Y)}) \right)^2 = \frac{\text{tr}^2(A)\text{tr}^2(B)}{(n-1)^2 \times \text{tr}(AA) \times \text{tr}(BB)}$$
\[ E_{\text{perm}} \left( RV_{(X,Y)} \right)^2 = \frac{1}{\text{trace} \left( AA \right) \text{trace} \left( BB \right)} E_{\text{perm}} \left( \text{trace} \left( AB \right) \right)^2 \]
\[ = \frac{1}{n! \times \text{trace} \left( AA \right) \text{trace} \left( BB \right)} \sum_{\sigma \in \Omega} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^\prime x_j^\prime x_i^\prime x_j^\prime y_{\sigma(i)}^\prime y_{\sigma(j)} \cdot y_{\sigma(i)} y_{\sigma(j)} \]
\[ + \sum_{i=1}^{n} \sum_{j=1}^{n} \left( y_{\sigma(i)}^\prime y_{\sigma(j)} \right)^2 \]
\[ + 2 \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1,j \neq k, i} \left( y_{\sigma(i)}^\prime y_{\sigma(j)} \right) \sum_{\sigma \in \Omega} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^\prime x_j^\prime x_i^\prime x_j^\prime y_{\sigma(i)}^\prime y_{\sigma(j)} \cdot y_{\sigma(i)} y_{\sigma(j)} \]
\[ + 2 \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1,j \neq k, i} \sum_{\sigma \in \Omega} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^\prime x_j^\prime x_i^\prime x_j^\prime y_{\sigma(i)}^\prime y_{\sigma(j)} \cdot y_{\sigma(i)} y_{\sigma(j)} \]
\[ + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1,j \neq k, i} \sum_{\sigma \in \Omega} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^\prime x_j^\prime x_i^\prime x_j^\prime y_{\sigma(i)}^\prime y_{\sigma(j)} \cdot y_{\sigma(i)} y_{\sigma(j)} \] (11)

According to Kazi-Aoual’s formula:

\[ V_{\text{per}, \text{trace}(AB)} = \frac{2((n-1) \times \text{trace}(AA) - (\text{trace}(A))^2) \times ((n-1) \times \text{trace}(BB) - (\text{trace}(B))^2)}{(n-1)^2(n+1)(n-2)} \]
\[ + \frac{(n(n+1) \sum_{i=1}^{n} [x_i^\prime x_i]^2 - (n-1)((\text{trace}(A))^2 + 2\text{trace}(AA))) \times (n(n+1) \sum_{i=1}^{n} [y_i^\prime y_i]^2 - (n-1)((\text{trace}(B))^2 + 2\text{trace}(BB)))}{(n+1)(n-1)(n-2)(n-3)} \] (12)

The variance of permutated \( RV_{(X,Y)} \) can be calculated as follows:

\[ V_{\text{per}, \text{RV}(XY)} = \frac{2((n-1) - (\text{trace}(A))^2)((n-1) - (\text{trace}(B))^2)}{(n-1)^2(n+1)(n-2)} \]
\[ + \frac{(n(n+1) \sum [x_i^\prime x_i]^2 - (n-1)((\text{trace}(A))^2 + 2\text{trace}(AA))) \times (n(n+1) \sum [y_i^\prime y_i]^2 - (n-1)((\text{trace}(B))^2 + 2\text{trace}(BB)))}{(n+1)(n-1)(n-2)(n-3)} \] (13)

Note that Equation 13 is not the exact value of variance of permutated RV, however, when \( n > 8 \), the formula is a very good approximation.

We define \( \gamma_{\text{perm}}(RV_{(X,Y)}) \) the skewness of the permutated \( RV_{(X,Y)} \), According to Kazi-Aoual’s expression, it can be calculated as follows:
\[ \gamma_{\text{perm}}(RV_{(X,Y)}) = \gamma_{\text{perm}}\left(\frac{\text{trace}(AB)}{\sqrt{\text{trace}(AA)\text{trace}(BB)}}\right) = \frac{1}{\sqrt{\text{trace}(AA)\text{trace}(BB)}} \gamma_{\text{perm}}(\text{trace}(AB)) \]

\[ = \frac{1}{\sqrt{\text{trace}(AA)\text{trace}(BB)}} \mu_z(\text{trace}(AB)) \left[ V_{\text{perm}}(RV_{(X,Y)}) \right]^{3/2} \]

\[ = \frac{E_{\text{perm}}(\text{trace}(AB)^3) - E_{\text{perm}}(\text{trace}(AB))V_{\text{perm}}(\text{trace}(AB)) - \left[ E_{\text{perm}}(\text{trace}(AB)) \right]^2}{\sqrt{\text{trace}(AA)\text{trace}(BB)}[V_{\text{perm}}(RV_{(X,Y)})]^{3/2}} \]

### 2.3 The temporal autocorrelation and the effective degree of freedom of the fMRI data

Due to various sources such as cardiac, respiratory, vasomotor sources, fMRI data exhibit strong time serial autocorrelation; the pre-processing steps of the spatial smoothing and temporal filtering also greatly increase the spatial and temporal correlation, which make the time point in the BOLD signal not statistically independent. Time serial correlations play an important role in assessing the significance of a test statistic, so the effective degree of freedom of the fMRI data under the null hypothesis must be estimated and corrected. Ignoring serial correlations will lead to too lenient and therefore invalid tests. Though we know that the effective degree of freedom is less than the number of the fMRI time acquisitions, the precise value of the effective degree of freedom can not be obtained due to various reasons. One ordinary way of solving this problem is using a rough correction factor. Experiments have found that the correction factor is associated with the number of the time acquisition, Here, we do not calculate the effective degree of freedom or the correction factor explicitly, we integrate the correction factor of the effective degree of freedom into the following approximated distribution.

### 2.4 The approximation of the fMRI data

Normal distribution is a good choice to approximate the permuted RV coefficient. For testing the significance of the permuted RV, only the first two moments is needed. The approximation of the permuted RV can be denoted as:

\[ \frac{RV(X,Y) - E_{\text{perm}}(RV_{(X,Y)})}{\sqrt{V_{\text{perm}}(RV_{(X,Y)})}} \sim N(0,1) \]

(15)

However, from the above calculation of the \( \gamma_{\text{perm}}(RV_{(X,Y)}) \), we know that the permutation distribution of the RV coefficient is remarkably skewed to the right, so we considered a transformation of the first two moments to obtain a better approximation which is more nearly normal. Meanwhile, as we have discussed in the above session, fMRI data exhibit strong time serial autocorrelation which will make the p-value too lenient and therefore invalid tests, and a correction factor must be considered to correct the time serial autocorrelation. Experiments also find that the variance of permuted RV decreases when the number of observations increases, the normalized RV-coefficient increases linearly with the numbers of observation increases. Combining the above two factors, we constructed a statistic which is the transformation of the first two moments of the permuted RV. This transformed statistic is considered to be a better approximation of the normal distribution:
Then the precise p-value can be obtained by calculating the above Z-score.

3. THE ACQUISITION OF FMRI DATA

Twelve healthy subjects (6 females, age: 20-27 years) with normal vision participated in this study after giving their informed consent. The study was approved by the research ethics committee at the Tiantan Hospital. The subjects performed a resting state scan first. Then they participated in a localizer task (e.g., Kanwisher et al., 1997; Tong et al., 2000) where they were instructed to pay close attention to a sequence of images of faces or common objects appearing on a screen. During the localizer period, participants were instructed to press a button when they saw two identical pictures appearing in a row to ascertain the participant to be attentive. The resting-state scan contained 180 time acquisitions, the localizer tasks contained 177 time acquisitions.

The MR imaging was carried out using a 3.0 T MR Scanner (Siemens Trio Tim). Functional images were collected axially using a T2*-weighted gradient-echo echo planar imaging (EPI) sequence (TR/TE=2000/30 ms; 32 slices, 4mm sickness; matrix=64×64) covering the whole brain with a resolution of 3.75×3.75mm. Structural images were acquired with a three-dimensional enhanced fast gradient-echo sequence with a thickness of 1mm and a resolution of 1×1mm.

4. THE ANALYSIS OF FMRI DATA

Preprocessing was performed on each participant’s scan data by using SPM5 software (www.fil.ion.ucl.ac.uk/spm/) and in house Matlab code. The first three volumes of each fMRI scan were discarded. Scans were slice timing corrected, spatially realigned, normalized into the standard MNI atlas space according to the segmented grey images; re-sampled to 2mm cube voxels.

For each participant’s localizer scan, a standard general linear model analysis was conducted with SPM5’s restricted maximum likelihood (ReML) estimation. Each participant’s face sensitive regions were identified by contrasting the face block with the common object block (threshold p=.0001). The most active regions were localized in the right lateral posterior fusiform gyrus, which is in line with the previous studies about the location of the fusiform face area (FFA) in occipital-temporal cortex. Then each individual’s FFA was precisely selected as regions of interest (ROI) for further analysis.

Each participant’s resting state scan were processed in the following manner: (1) global proportional scaling was performed to yield whole brain intensity value of 1000; (2) drifts were de-trended by second-order polynomial detrending; (3) six parameters of head motion, signals in ventricular regions and signals extracted from white matter were removed by linear regression; (4) discrete cosine transform domain (DCT) band-pass filter (0.01-0.08Hz) was performed to retain the low frequency signals only.

For each participant’s resting scan data, the time signals for all the voxels in the FFA were extracted and considered as region seed to map a whole brain RV-coefficient based functional connectivity. RV-coefficient was used here to measure
the similarity of ROI and each local region of whole brain. To obtain a continuous map, we move a cube which comprises 27 voxels through the measured volume; the cube is centered on each voxel in the brain volume in turn, then the RV-coefficient between the time signals in each subject’s FFA and the time signals in each 27 neighboring local region is calculated.

We use equation 16 to transform the value of each voxel’s RV-coefficient to its corresponding Z-score, so we can use the Z criterion of 1.69 (Threshold: $p=0.0077$) to test the significance of each voxel’s RV-coefficient under the null hypothesis in functional connectivity map. For comparison purpose, we also map the functional connectivity with the Pearson correlation coefficient.

5. RESULT

We illustrate the sampling distributions of the RV-coefficients, the normalized RV-coefficients and the transformed RV-coefficients. The sampling data sets containing 206479 samples were extracted from one subject’s functional connectivity maps. As can be seen in Fig.1, the value of RV-coefficients is distributed between 0 and 1, the distribution of the normalized RV-coefficients is similar to a normal distribution, however, it is overestimated severely. The distribution of RV and normalized RV were skewed greatly to the right. While our transformed method showed better approximation to the normal distribution. Fig.2 gave the maps of the functional connectivity in resting state (selecting each this subject’s fusiform face area as regions of interest) and its significance test ($p=0.0077$, uncorrected). From Fig.2 we can see that this subject’s right fusiform face areas (FFA) has significant connectivity with most of the occipital-temporal cortex, especially significant in parts of the primary visual cortex, such as left middle fusiform gyrus, bilateral middle occipital gyrus, middle and inferior frontal gyrus and part of the parietal lobe. This map also shows the fine-scale patterns of the functional connectivity between right FFA and the most complicated occipital-temporal regions.

6. CONCLUSIONS

The statistic of RV-coefficient was used in fMRI to extract the fine-scale patterns of the functional connectivity of brain. To test the significant of RV-coefficient, A new strategy was proposed in this article which made the significance test of large observations of RV possible and precise, this strategy can greatly reduce the computational complexity and avoid “calculation catastrophe”. We then use the statistic of RV to extract the map of functional connectivity from fMRI and test its significance with the strategy proposed here. The experiment result shows this strategy got more close approximation and achieved better performance.

7. ACKNOWLEDGMENTS

This paper is supported by the Project for the National Basic Research Program of China (973) under Grant No.2006CB705700, Changjiang Scholars and Innovative Research Team in University under Grant No.IRT0645, CAS Hundred Talents Program, CAS scientific research equipment develop program (YZ0642, YZ200766), 863 program under Grant No. 2006AA04Z216, the Joint Research Fund for Overseas Chinese Young Scholars under Grant No.30528027, the National Natural Science Foundation of China under Grant No. 30672690, 30600151, 30500131, 60532050, 60621001, Beijing Natural Science Fund under Grant No. 4071003, and NIH R01 HD046526.
Fig. 1. The sampling distributions of the RV-coefficients, the normalized RV-coefficients and the transformed RV-coefficients. The sampling data sets containing 206479 samples were extracted from one subject’s functional connectivity maps.

Fig. 2. Map of RV-based functional connectivity in resting states by using FFA as region of interests and its significance test. (a) Values of RV-coefficient from single slice. (b) Map of transformed RV at p<0.0077 (Z-score>1.69)

REFERENCES


