Numerical study of resting-state fMRI based on kernel ICA

Zhu Dongjuan   Wang Xunheng   Ruan Zongcai

(Research Center for Learning Science, Southeast University, Nanjing 210096, China)

Abstract: In order to facilitate the extraction of the default mode network(DMN), reduce the data complexity of the functional magnetic resonance imaging (fMRI) and overcome the restriction of the linearity of the mixing process encountered with the independent component analysis (ICA), a framework of dimensionality reduction and nonlinear transformation is proposed. First, the principal component analysis (PCA) is applied to reduce the time dimension $153\ 594 \times 128$ of the fMRI data to $153\ 594 \times 5$ for simplifying complexity computation and obtaining 95% of the information. Secondly, a new kernel-based nonlinear ICA method referred as the kernel ICA(KICA) based on the Gaussian kernel is introduced to analyze the resting-state fMRI data and extract the DMN. Experimental results show that the KICA provides a better performance for the resting-state fMRI data analysis compared with the classical ICA. Furthermore, the DMN is accurately extracted and the noise is reduced.

Key words: kernel independent component analysis; principal component analysis; functional magnetic resonance imaging (fMRI); resting-state

Functional magnetic resonance imaging(fMRI) based on a blood oxygen level dependent (BOLD) signal is a noninvasive medical imaging technique. Recently, it is widely used in the area of imaging for its high temporal and spatial resolution.

Resting-state fMRI is defined as no cognitive tasks during an fMRI scan. The experiment is accomplished under the resting-state. In other words, participants should simply remain motionless, keep alert with their eyes closed and do not concentrate on anything specifically[1]. Since the resting-state is the most complex state of the human brain, it has been a hot topic in researches on the resting-state. Biswal et al. [2–3] found slow fluctuations of BLOD signals in the human brain when a human is under the resting-state. Gusnard et al. [4–5] proposed the concept of the default mode network(DMN). The DMN is formed by regions and shows coherent functional behavior while resting[6–8]. A great deal of evidence suggests that the default mode network is a special mode. It becomes significant under the resting-state and decreases when individuals focus on external inputs, but it does not disappear.

The independent component analysis (ICA) which solves problems similar to "the cocktail party" problem[9] is popular in many fields, such as image processing and face recog-

nition etc. However, the ICA is carried out with the restriction of the linearity of the mixing process. For solving this problem, the kernel ICA(KICA) is utilized to overcome the restriction of the linearity of the mixing process. In principle, the KICA transforms the original data into a higher feature space and analyzes projection data through "kernel trick"[10].

1 Kernel ICA Plus PCA

Before the application of the ICA, the principle component analysis (PCA) is performed to reduce the dimension of the data first. The reasons why the PCA should be first used are interpreted in two parts: 1) As is known, the fMRI data is great, especially the spatial dimension, and data reduction should be used in order to make it easier to carry out the ICA; 2) It is assumed that independent sources are fewer than time points (spatial ICA) or spatial voxels (time ICA)[11].

1. 1   PCA

The PCA is a typical method of feature extraction and data compression. After SPM(www. fil. ion. ucl. ac. uk _ spm) pre-processing, the spatial dimension of the fMRI data will reach $53 \times 63 \times 46$. However, if a high dimensional data is directly used on the ICA, it will cost too much computation time. For simplicity of computation, the PCA is utilized to reduce the dimension of the times points.

The voxels of each subject can be rearranged into a vector $X_i(1 \times N)$ with t time orders. So we obtain

$$X = [\ X_1 \quad X_2 \quad \dots \quad X_t\ ]^{\mathrm{T}} \tag{1}$$

Before the process of the PCA, the data is centered, and

$$\bar{X} = [\ X_1 - \pmb{\mu} \quad X_2 - \pmb{\mu} \quad \dots \quad X_t - \pmb{\mu}\ ]^{\mathrm{T}} \tag{2}$$

In Eq. (2), $\pmb{\mu}$ is the mean vector of X and is defined as

$$\pmb{\mu} = \frac{1}{t} \sum_{i=1}^{t} X_i \tag{3}$$

The covariance matrix of X is estimated as

$$C = (X - \bar{X})(X - \bar{X})^{\mathrm{T}} \tag{4}$$

Through eigenanalysis, the eigenvectors $\pmb{v}_k$ and the corresponding eigenvalues $\pmb{d}_k$ can be obtained. Sorting the eigenvalues d_1, d_2, …, d_t according to a descending order, the eigenvectors are also sorted according to the eigenvalues.

Finally, due to the eigenvalues d_1, d_2, …, d_t falling sharply, a threshold value is set to determine the number of components. Assuming that there are n eigenvalues over the threshold, the first $n(n < t)$ eigenvector which contains the most information is chosen, and then the data is projected

into the feature space. Thus, the data is

$$\hat{X} = V^{\mathrm{T}} X \qquad (5)$$

where $\hat{X}$ contains the main information of the fMRI data.

1. 2   Kernel ICA

The ICA, a statistical method, is utilized to discover blind factors(sources or features) from a set of measurements or observed data[12]. A typical ICA model assumes that the source signals are statistically independent and non-Gaussian, with an unknown but linear mixing process[11].

The observed signal is defined as

$$X = [\, X_1 \quad X_2 \quad \ldots \quad X_m \,]^{\mathrm{T}}$$

which stands for the input of the model, where m is the number of signals. And the output is a set of source signals which can be described as

$$S = [\, S_1 \quad S_2 \quad \ldots \quad S_l \,]^{\mathrm{T}}$$

where l is the number of sources. According to the idea of the ICA, the relationship between the input and output signals can be represented as

$$X = AS \qquad (6)$$

$$S = WX \qquad (7)$$

where A is a mixing matrix with $m \times l(m \geqslant l)$, and W is an unmixing matrix with $l \times m$. It is noted that A and W are reciprocal. Thus, the goal of the ICA is to find W and make it as close to A^{-1} as possible.

However, it is easy to find that the ICA is only adapted to the linearity of the mixing process. Practically, the real data set is always nonlinear as far as their dimensionality is concerned. Fortunately, nonlinear methods, such as the kernel method and the manifold method, are proposed to deal with those above-mentioned situations. In the nonlinear methods, the kernel method is popular in the statistical method, such as the PCA and the linear discriminant analysis(LDA), for its robustness and effectiveness. In the literature, kernel methods are used to expand the linear eigenspace into nonlinear eigenspace for face recognition, or expand a linear classifier to a nonlinear situation. Owing to its success in pattern recognition, the kernel method is used to expand the ICA for nonlinear mixing processing in this paper.

Regarding the function of kernel methods, kernel functions are satisfied with the mercer theorem. Generally, the Gaussian kernel, the polynomial kernel and the Hermite kernel are used in different applications. In this paper, the Gaussian kernel is applied and the KICA, based on the kernel generalized variance(KGV) is employed.

The mixing signal we observed is denoted as

$$X = [\, X_1 \quad X_2 \quad \ldots \quad X_M \,]$$

and the estimated source signals through the KICA can be written as

$$S = [\, S_1 \quad S_2 \quad \ldots \quad S_M \,]$$

where M is the number of voxels. Similar to the ICA, the goal of the KICA is also found to be an unmixing matrix, which is defined as W, for making blind signals separation.

Thus, the relationship between blind signals and estimated signals can be expressed as

$$S_j = WX_j \qquad j = 1, 2, \ldots, M$$

where $S_j \in \mathbf{R}^{n \times 1}$, and n is the number of independents. We can obtain n Gram matrices K_1, K_2, ..., K_n, and K_i can be computed from S_{ij}, $j = 1, 2, \ldots, M$. Based on the Gram matrices, matrix $K_k{}^{[10]}$ can be denoted as

$$K_k = \begin{bmatrix} \left(K_1 + \dfrac{Mk}{2}I\right)^2 & K_1K_2 & \ldots & K_1K_n \\ K_2K_1 & \left(K_2 + \dfrac{Mk}{2}I\right)^2 & \ldots & K_2K_n \\ \vdots & \vdots & & \vdots \\ K_nK_1 & K_nK_2 & \ldots & \left(K_n + \dfrac{Mk}{2}I\right)^2 \end{bmatrix} \qquad (8)$$

and D_k is denoted as a block-diagonal matrix with diagonal elements of K_k. In Eq. (8), k is an adjustment factor.

$$\hat{\delta}_F^k(K_1, \ldots, K_n) = \frac{\det K_k}{\det D_k} \qquad (9)$$

By the kernel generalized variance defined in Eq. (9), we can obtain the contrast function as follows:

$$\hat{I}_{\delta F}(K_1, \ldots, K_n) = -\frac{1}{2}\log \hat{\delta}_F^k(K_1, \ldots, K_n) \qquad (10)$$

Finally, the KICA algorithm can be concluded as below:
Observed data vectors: $X = [\, X_1 \quad X_2 \quad \ldots \quad X_M \,]$;

Kernel: $K(x, y) = \exp\left(\dfrac{1}{2\sigma^2} \| x - y \|^2\right)$;

1) Whiten the data;
2) Initialize the unmixing matrix W;
3) Compute the centered Gram matrices K_1, K_2, ..., K_n of the estimated sources $S = [\, S_1 \quad S_2 \quad \ldots \quad S_M \,]$;

4) Define contrast function as $C(W) = -\dfrac{1}{2}\log \hat{\delta}_{\delta F}(K_1, \ldots, K_n)$;

5) Minimize the value of $C(W)$;
Output: W.

2   Experiment and Discussions

The fMRI data is preprocessed by using the software package SPM to perform "slice timing", "realign", "spatial normalize" and "smooth", resulting in $53 \times 63 \times 46$ voxels. We select the middle slice(the 23rd slice) to estimate the source signals, and the results are compared with other algorithms.

In order to compare the results of the DMN extracted by the KICA, GroupICA(http: //icatb. sourceforge. net/gift/ gift_startup. php) for fMRI Toolbox is utilized to produce the reference result. The result of GroupICA is shown in Fig. 1. It is easily found that the first independent component (IC1) is the DMN. Then, we apply the KICA to ex-

tract the independent components from the fMRI data. Fig. 2 presents our analytical result of five independent components. We can see that the fifth independent component (IC5) in Fig. 2 is the DMN. Comparing Fig. 2 with Fig. 1, the components of the DMN through the KICA and GroupICA are similar. We can see that the noise level in Fig. 2 is much lower than that in Fig. 1. It implies that the KICA works much better than the classical ICA in the field of resting-state fMRI data analysis.
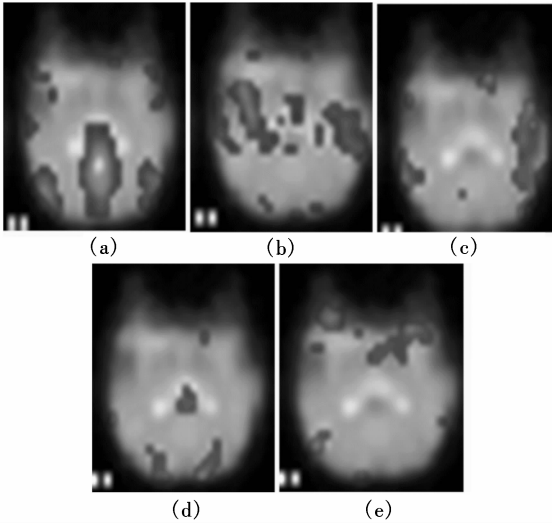
Fig. 1 Independent components decomposed by GroupICA. (a) IC1; (b) IC2; (c) IC3; (d) IC4; (e) IC5
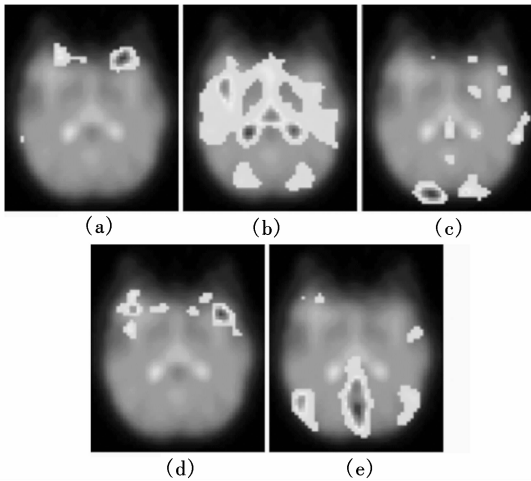

Fig. 2 Independent components decomposed by KICA. (a) IC1; (b) IC2; (c) IC3; (d) IC4; (e) IC5

The DMN extracted by the KICA in different slices is presented in Fig. 3. The results confirm that the KICA can be used to analyze the resting-state fMRI data and it can work well.

3 Conclusion

In this paper, a novel method for extracting the resting-state network is proposed. Compared with GroupICA, the KICA is more effective in analyzing resting-state fMRI data. By the comparison of extracting default mode network in the same slice, we find that the KICA is more effective than GroupICA. The KICA can be applied to analyze resting-

state fMRI data. Moreover, it has great potential for its fine flexibility and robustness.
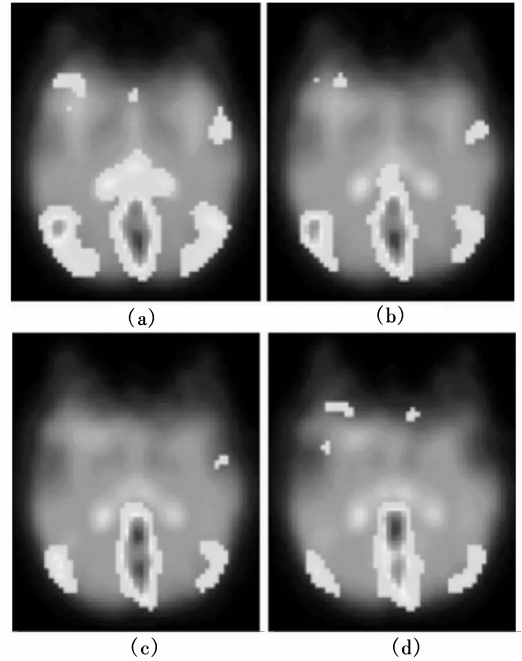
Fig. 3 Default mode network extracted from different slices by KICA. (a) The 22nd slice; (b) The 23rd slice; (c) The 24th slice; (d) The 25th slice

References

[1] Wang Xunheng, Zhou Zhenyu, Jiao Yun, et al. Investigating the structure of default mode network with social network analysis [C]//*Human Brain Mapping*. San Francisco, CA, USA, 2009: 101.

[2] Biswal B, Yetkin F Z, Haughton V M, et al. Functional connectivity in the motor cortex of resting human brain using echo-planar mri [J]. *Magnetic Resonance in Medicine*, 1995, **34**(4): 537 − 541.

[3] Damoiseaux J S, Rombouts S A R B, Barkhof F, et al. Consistent resting-state networks across healthy subjects [J]. *Proceedings of the National Academy of Sciences*, 2006, **103**(37): 13848 − 13858.

[4] Gusnard D A, Raichle M E. Searching for a baseline: functional imaging and the resting human brain [J]. *Nature Review Neuroscience*, 2001, **2**(10): 685 − 694.

[5] Raichle M E, MacLeod A M, Snyder A Z, et al. A default mode of brain function [J]. *Proceedings of the National Academy of Sciences*, 2001, **98**(2): 676 − 682.

[6] Fair D A, Cohen A L, Dosenbach N U F, et al. The maturing architecture of the brain's default network [J]. *Proceedings of the National Academy of Sciences*, 2008, **105**(10): 4028 − 4032.

[7] Greicius M D, Srivastava G, Reiss A L, et al. Default-mode network activity distinguishes alzheimer's disease from healthy aging: evidence from functional MRI [J]. *Proceedings of the National Academy of Sciences*, 2004, **101**(13): 4637 − 4642.

[8] Buckner R L, Andrews-Hanna J R, Schacter D L. The brain's default network: anatomy, function, and relevance to disease [J]. *Annals of the New York Academy of Sciences*, 2008, **1124**(1): 1 − 38.

[9] Bell A J, Sejnowski T J. An information maximization approach to blind separation and blind deconvolution [J]. *Neural Computation*, 1995, **7**(6): 1129 − 1159.

[10] Bach F R, Jordan M I. Kernel independent component anal-

ysis [J]. *Journal of Machine Learning Research*, 2002, **3**(3): 1－48.

[11] Calhoun V D, Adali T, Pearlson G D, et al. Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms [J]. *Human Brain Mapping*, 2001, **13**(1): 43－53.

[12] Calhoun V D, Adali T. Unmixing fMRI with independent component analysis [J]. *Engineering in Medicine and Biology Magazine*, 2006, **25**(2): 79－90.

基于核独立成分分析的静息态 fMRI 数据研究

朱冬娟　　王训恒　　阮宗才

（东南大学学习科学研究中心，南京 210096）

摘要：为了方便提取静息态默认网络，降低功能核磁共振（fMRI）数据复杂度，克服独立成分分析只适合于源信号线性混合的限制，提出了特征降维和非线性变换的框架. 首先采用主成分分析对 fMRI 信号的时间维度进行降维，将原始维度为 153 594×128 的 fMRI 数据降至 153 594×5，以达到降低计算复杂度的目的，并保留 95% 的信息成分. 然后利用基于高斯核的非线性独立成分分析即核独立成分分析来分析静息态 fMRI 数据并提取默认网络. 实验结果表明，在分析静息态 fMRI 数据的过程中，核独立成分分析不仅能准确提取默认网络，而且降低了噪声，所得到的结果优于普通独立成分分析.

关键词：核独立成分分析；主成分分析；功能核磁共振；静息态

中图分类号：R318.04