Recent advances in the data analysis method of functional magnetic resonance imaging and its applications in neuroimaging*

TIAN Jie[1,2**], YANG Lei[1] and HU Jin[1]

(1. Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China; 2. Graduate School of Chinese Academy of Sciences, Beijing 100039, China)

Received November 28, 2005; revised February 20, 2006

Abstract Functional magnetic resonance imaging (fMRI) has opened a new area to explore the human brain. The fMRI can reveal the deep insights of spatial and temporal changes underlying a broad range of brain function, such as motor, vision, memory and emotion, all of which are helpful in the clinical investigation. In this paper, we introduce some recent-developed algorithms for fMRI signal detection such as model-driven method (general linear model, deconvolution model, non-linear model, etc.) and data-driven method (principle component analysis, independent component analysis, self-organization mapping, clustered constrained non-negative matrix factorization, etc.). We also propose several important applications of neuroimaging and point out their shortcomings and future perspectives.

Keywords: fMRI, data analysis, neuroimaging.

The idea that regional cerebral blood flow (CBF) can reflect neuronal activity was confirmed by the experiments of Roy and Sherrington in 1890[1]. This concept is the basis of all the hemodynamic-based brain imaging techniques being used today. The focal increase of CBF is directly related to the neuronal activity. Thus, CBF changes have been used to measure the functions of brain. CBF changes can be detected by fMRI based on the fact that the MRI signal is sensitive to many hemodynamic parameters (such as blood flow, blood volume and oxygenation), which was reported in the early 1990s by a number of researches[2—4], Currently, fMRI is the most influential noninvasive technology with venous blood oxygenation level-dependent (BOLD) magnetic resonance imaging (MRI) contrast. BOLD fMRI signal advances with the increasing local neuronal activity. The neuron activation takes oxygen and glucose from the surrounding capillary. At first, deoxyhaemoglobin (dHb) concentrates at the site of activation due to increased consumption of oxygen and glucose. dHb is paramagnetic and leads to weaker BOLD fMRI signal. With the persistence of neural activity, the cerebral blood flow overcompensates the demand of oxygen. Consequently, the haemoglobin (Hb) concentrates, which is diamagnetic and makes stronger fMRI signal. Details of this process can be described by certain mathematic models.

In a typical fMRI investigation, the subject's brain function is activated by performing certain task in a MR scanner (for example, to make finger tap or to see some images while scanning). The whole session contains mainly four steps experimental design, image acquisition, preprocessing (usually including image registration and noise removal), and signal detection. Several types of signals can be encoded within the hemodynamic signals measured by fMRI (for example, task-related components, neurophysiological components, and noise). And the intensity difference of MR signal between the task and rest is only 2%—5% of the local image. After preprocessing of fMRI data, signal detection corresponds to modelling the data to partition observed neurophysiological responses into the components of interest, confounds or components of no interest and error terms. And the research of fMRI data analysis aims to separate the components of interests. Our work is focused on the discussion of signal detection which is the core of data analysis.

The methods are usually classified into two categories: model-driven and data-driven methods. In the model-driven category, general linear model (GLM) is the most popular one. It is a framework that in-

* Supported by National Science Fund for Distinguished Young Scholars of China (Grant No. 60225008), National High Technology Research and Development Program of China (Grant No. 2004AA420060), National Natural Science Foundation of China (Grant Nos. 30370418, 90209008, 60302016, 30270403), and Beijing Natural Science Fund (Grant No. 4051002,4042024)

** To whom correspondence should be addressed. E-mail: tian@doctor.com

cludes simple t-test, analysis of variance (ANOVA) and multiple regressions. Upon the completion of an experimental design, the design matrix of GLM can be specified and the classic parameter estimation methods (such as multiple regression) will be applied. A parameter map is obtained as a consequence and its statistic parameter inference is made under the null-hypothesis that the voxel is nonactivated. The most popular free software Statistical Parametric Mapping (SPM) has taken GLM as its solution. Deconvolution is another often-used model-driven method[5,6]. At first the impulse response function (IRF) is estimated and then convolves the IRF with the stimulus paradigm to yield the estimated response. Then, various statistics are calculated to indicate the "goodness" of the fit. Finally, according to a predefined level of statistics, the activated voxels can be identified. Analysis of functional neuroimages (AFNI) is implemented through the deconvolution model. Fast-orthogonal search (FOS) is a kind of nonlinear method. FOS provides a more accurate model description and is an efficient algorithm which achieves real time computation. Obviously, model-driven method seriously depends on the prior experimental design knowledge. However, when there is only a mass of data without any prior knowledge about experiments, the data-driven method provides an option to analyze the brain activity by means of independent component analysis (ICA), principle component analysis (PCA), etc. These two methods both search for the components which can explain the measured fMRI signal "well". The only difference between the two methods is that the principal components are orthogonal and the independent components are statistically independent. Cluster analysis is another data-driven method which includes c-means, fuzzy cluster analysis (FCA), self-organizing mapping (SOM), and so on. These methods are applied to hunt for a class of points with similar properties. Among them, clustered constrained non-negative matrix factorization is a relatively new method to preserve the intrinsic geometric characteristics of the data.

There is no such a review which summarizes all the existing methods for data analysis. And no predominant advantages have been found in either kind of methods. They are complementary.

fMRI is used not only in cognitive research, but also in clinical applications such as the navigation. The ultimate success of these research is based on the accuracy model of the relationships between the measured signal and the neuronal activity. In this review, we first briefly explain the fMRI imaging principles and then discuss some standard methods for characterization of data model.

1 fMRI data analysis

This section describes the analysis steps of fMRI data including both the preprocessing and signal detection. The preprocessing usually involves two steps. First, the image of time series must be realigned and then registered to a standard coordinate space. Subject motion during a fMRI examination results in a significant loss of functional information. To overcome this side effect, all the slices must be realigned. Under the assumption of the rigid motion, realignment only requires a rigid transformation for each image, e.g. aligning every image to the desired image of the time series. To compare time series among different subjects, it is necessary to map the anatomical coordinates into a standard coordinates space for each subject. The Talariach space and Montreal Neurological Institute (MNI) space are the often-used ones. In this case, the images must be spatially normalized which require nonlinear transformations[7]. Secondly, noise removal is a routine for fMRI data preprocessing for two reasons[8,9]: one is the increase of the signal-to-noise ratio (SNR), the other is the interpretation of the images as gaussian random fields. Noise removal may also be called spatial smoothing and can be done by convolving the image with a lowpass filter kernel. After denoising, the images are ready for signal detection.

1.1 Model-driven methods

In this section, we first review some linear model-driven fMRI data analysis methods such as general linear model and deconvolution model. For a comprehensive description of the hemodynamic model, nonlinear model must be placed then.

1.1.1 General linear model GLM stands for general linear model and its root may go back to the origins of mathematical thought—the theory of algebraic invariants in the 1800s. The theory was developed by several great mathematicians such as Gauss (he first established the fundamental theorem of algebra which states that any complex polynomial must have a complex root), Boole (who initiated the "algebraic invariants"), Cayler (the founder of the polyno-

mial discriminant which is originally known as hyperdeterminants) and Sylvester (the inventer of reciprocants-differential invariants theory). The algebraic invariants theory seeks to identify those quantities in systems of equations which remain unchanged with the linear transformation such as eigenvalues, eigenvectors, determinants and the correlation between two variables.

Friston[10—13] first introduced the concept of GLM into fMRI and nowadays it has been applied in many areas. GLM is a voxel by voxel method in fMRI application. It models the data at each voxel as a linear combination of explanatory variables plus a residual error term[13]. It then creates a mapping to reflect the statistical significance and characterize the specific regional response. GLM in fMRI is a fairly mature framework and has been implemented as a free software called statistical parametric mapping (SPM). GLM is at the center of the frame. Here, we sum up the steps of the model.

Let us denote by $\boldsymbol{X}$ ($N \times T$ matrix) one session of a fMRI data set, where N is the number of voxels in the data set and T is the length of the time series. Thus, $X_n(t)$ is the signal at voxel n and time t. The model assumes that 1) the response to the stimulation is linear time-invariant; 2) canonical hemodynamic response function (hrf) is assumed to be identical for all the voxels; 3) the noise ε_n is independent and identically distributed normal random variables with zero mean and variance σ^2, written as $\varepsilon_n \overset{\text{iid}}{\sim} N(0,\sigma^2)$. Then the most basic signal model is

$$X_n(t) = b_0 + \sum_{m=1}^{M} b_m(n) h P_m(t) + \varepsilon_n(t), \tag{1}$$

where $P_m(t)$, $t = 1,2,\cdots,T$, $m = 1,2,\cdots,M$ is the time course of the effect and h is the canonical hrf. $b_m(n)$ is the amplitude of the stimulus responses and $\varepsilon_n(t)$ is the noise term. Consequently, we assume the signal is centered, and then $b_0 = 0$.

Let us denote $g_m(t) = h * P_m(t)$, $m = 1,2,\cdots,M$. Then we obtain a matrix form with the design matrix $\boldsymbol{G} = \{g_m(t)\}$:

$$\boldsymbol{X} = \boldsymbol{G}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{2}$$

The design matrix $\boldsymbol{G}$ contains the explanatory variables related to the specific experimental conditions and each column is associated with an unknown parameter in vectors $\boldsymbol{\beta}$. If and only if the design matrix $\boldsymbol{G}$ is full rank, least square estimate of $\boldsymbol{\beta}$ is uniquely given by

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G})^{-1}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{X}. \tag{3}$$

Through the Gauss-Markov theorem, the least square estimation is the maximum likelihood one when the errors are normally distributed and it is also the best linear unbiased estimation. That means that in all the linear estimations whose expectation is the true value, the least square estimate has the minimum variance.

The residual variance σ^2 is estimated by the residual mean square:

$$\hat{\sigma}^2 = \frac{\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\varepsilon}}{T-p} \sim \sigma^2 \frac{\chi^2_{T-p}}{T-p}, \tag{4}$$

where $p = \mathrm{rank}(\boldsymbol{G})$. The least square estimation is normally distributed: $\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2(\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G})^{-1})$, then for a column vector $\boldsymbol{c}$ containing L weights,

$$\boldsymbol{c}^{\mathrm{T}}\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{\beta}, \sigma^2\, \boldsymbol{c}^{\mathrm{T}}(G^{\mathrm{T}}G)^{-1}\boldsymbol{c}). \tag{5}$$

Then $\boldsymbol{c}^{\mathrm{T}}\hat{\boldsymbol{\beta}}$ can be assessed by using

$$\frac{\boldsymbol{c}^{\mathrm{T}}\hat{\boldsymbol{\beta}} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{\beta}}{\sqrt{\hat{\sigma}^2\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G})^{-1}\boldsymbol{c}}} \sim t_{T-p}, \tag{6}$$

where t_{T-p} is the student distribution with the freedom $T-p$. In SPM, all tested null hypotheses are of the form $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{\beta} = 0$ and all the SPM tests based on T-distribution are one side.

We suppose a model with parameter vector $\boldsymbol{\beta}$ which can be partitioned into two parts: $\boldsymbol{\beta}^{\mathrm{T}} = [\boldsymbol{\beta}_1^{\mathrm{T}} \square \boldsymbol{\beta}_2^{\mathrm{T}}]^{\mathrm{T}}$ and wish to test the hypothesis $\mathcal{H}$: $\boldsymbol{\beta}_1 = 0$. The corresponding partition of the design matrix $\boldsymbol{G}$ is $\boldsymbol{G} = [\boldsymbol{G}_1 \square \boldsymbol{G}_2]$, then the full model is

$$\boldsymbol{X} = [\boldsymbol{G}_1 \square \boldsymbol{G}_2] \begin{bmatrix} \boldsymbol{\beta}_1 \\ \cdots \\ \boldsymbol{\beta}_2 \end{bmatrix} + \boldsymbol{\varepsilon}. \tag{7}$$

When $\mathcal{H}$ is true, it becomes a reduced model: $\boldsymbol{X} = \boldsymbol{G}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$. Suppose the residual sums of squares for the full and reduced models are $s(\boldsymbol{\beta})$ and $s(\boldsymbol{\beta}_2)$ respectively, then extra sum of squares due to $\boldsymbol{\beta}_1$ after $\boldsymbol{\beta}_2$ is $s(\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2) = s(\boldsymbol{\beta}_2) - s(\boldsymbol{\beta})$. Under the assumption that $\mathcal{H}$ is true, $s(\boldsymbol{\beta})$ and $s(\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2)$ are independent and $s(\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2) \sim \sigma^2\chi^2_p$, where the degree of freedom is $p = \mathrm{rank}(\boldsymbol{G}) - \mathrm{rank}(\boldsymbol{G}_2)$. Then the hypothesis can be assessed by

$$F = \frac{(s(\boldsymbol{\beta}_2) - s(\boldsymbol{\beta}))/(p - p_2)}{s(\boldsymbol{\beta})/(T-p)} \sim F_{p-p_2,T-p}, \tag{8}$$

where $p = \mathrm{rank}(\boldsymbol{G})$, $p_2 = \mathrm{rank}(\boldsymbol{G}_2)$ and T is the total scan numbers.

In addition to standard GLM in SPM, we will discuss another popular model-driven method: deconvolution model. In this model, the impulse response function (IRF) is estimated first, and the response is obtained by convolving the IRF with the stimulus paradigm. Then, various statistics are calculated to demonstrate the "goodness" of fit. Finally, the activated voxels can be identified by a predefined level of statistics.

1.1.2 Deconvolution model In the deconvolution model, the system is assumed to be linear time-invariant. For an arbitrary input, the response of a linear time-invariant system can be determined from its response to an impulse function (Dirac delta function). The impulse function $\delta(t)$ is a theoretical mathematic concept which has an infinite height but zero width

$$\delta(t) = \begin{cases} +\infty, & t = 0 \\ 0, & t \neq 0 \end{cases} \tag{9}$$

with the property that the area under its curve is unity:

$$\int_{-\varepsilon}^{+\varepsilon} \delta(\tau)\mathrm{d}\tau = 1, \quad \forall \varepsilon > 0. \tag{10}$$

Let us denote $f(t)$ as follows:

$$f(t) = \int_{-\infty}^{+\infty} f(\tau)\delta(t-\tau)\mathrm{d}\tau, \tag{11}$$

and the discrete form is

$$f(t) = \lim_{\Delta t \to 0} \sum_{n=-\infty}^{+\infty} f(n\Delta t)\delta(t - n\Delta t)\Delta t. \tag{12}$$

For the impulse function, we denote $h(t)$ as its corresponding output function which is also called impulse response function (IRF):

$$h(t) = \mathbb{T}(\delta(t)). \tag{13}$$

For an arbitrary input $f(t)$, assume that $y(t)$ is the corresponding output

$$y(t) = \mathbb{T}(f(t)). \tag{14}$$

Since the system is linear, the formulation can be written as

$$y(t) = \lim_{\Delta t \to 0} \sum_{n=-\infty}^{+\infty} f(n\Delta t)\mathbb{T}(\delta(t - n\Delta t))\Delta t. \tag{15}$$

And the system is time-invariant, then

$$y(t) = \lim_{\Delta t \to 0} \sum_{n=-\infty}^{+\infty} f(n\Delta t)h(t - n\Delta t)\Delta t. \tag{16}$$

In an integration form, we have

$$y(t) = \int_{-\infty}^{+\infty} f(\tau)h(t-\tau)\mathrm{d}\tau. \tag{17}$$

And using the convolution integral concept, we have

$$\int_{-\infty}^{+\infty} f(\tau)h(t-\tau)\mathrm{d}\tau \equiv f(t) \otimes h(t). \tag{18}$$

In fact, all the physical system is causal, which means the output at time t_0 is only determined by the inputs at times $t \leqslant t_0$. And by assuming the input is zero when $t < 0$, the convolution integral can be written as

$$y(t) = \int_0^t f(\tau)h(t-\tau)\mathrm{d}\tau. \tag{19}$$

In the discrete form, $y(t)$ can be

$$y(n\Delta t) = \sum_{m=0}^{n} f(m\Delta t)h(n\Delta t - m\Delta t)\Delta t. \tag{20}$$

For a numeric solution, we have

$$y_n = \sum_{m=0}^{p} f_{n-m}h_m, \quad n \geqslant p. \tag{21}$$

Of course, for a practical system, the measurement must have some noise, then denote by Z_n the real measurements and

$$Z_n = \sum_{m=0}^{p} f_{n-m}h_m + \varepsilon_n, \quad n \geqslant p, \tag{22}$$

where the noise term ε_n is often assumed to have the property: $\varepsilon_n \overset{\text{iid}}{\sim} N(0, \sigma^2)$.

For fMRI data set, we have the following matrix form:

$$\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{23}$$

where

$$\boldsymbol{Z} = \begin{bmatrix} Z_p \\ Z_{p+1} \\ \square \\ Z_{N-1} \end{bmatrix},$$

$$\boldsymbol{X} = \begin{bmatrix} 1 & p & f_p & \cdots & f_0 \\ 1 & p+1 & f_{p+1} & \cdots & f_1 \\ \square & \square & \square & \square & \square \\ 1 & N-1 & f_{N-1} & \cdots & f_{N-p-1} \end{bmatrix},$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ h_0 \\ \square \\ h_p \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_p \\ \varepsilon_{p+1} \\ \square \\ \varepsilon_{N-1} \end{bmatrix}.$$

Then the least square estimate of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Z}. \tag{24}$$

Now various statistic tests (such as t-test and F-test) can be applied to access the hypothesis. And according to a given threshold, the activated voxel can be identified.

1.1.3 Nonlinear model All the above mentioned methods are based on the assumption of linear time-invariant (LTI) system. However, LTI sytems may not be able to adequately model the dynamics of fMRI signal. Great effort has been made to develop the nonlinear methods for modelling the fMRI signal dynamics[14−16]. Recently, Li et al.[17] proposed a fast-orthogonal search (FOS) method to model the nonlinear dynamics of the fMRI signal based on the Volterra series. The authors first approximated a nonlinear system by the discrete time Volterra series as follows:

$$y(n) = h_1 + \sum_{j=1}^{R-1} h(j)x(n-j) + \sum_{j_1=1}^{R-1}\sum_{j_2=1}^{R-1} h(j_1,j_2)x(n-j_1)x(n-j_2), \tag{25}$$

where $h_1, h(j), h(j_1,j_2)$ represent the zero-order, first-order and second-order Volterra kernels, respectively, and they can be estimated from the knowledge of the system input $x(n)$ and output $y(n)$. R is the number of basis function. Eq. (25) can be briefly written as

$$y(n) = \sum_{m=1}^{M-1} a_m P_m(n) + e(n), \tag{26}$$

where $n = 1,2,\cdots,N$, and N is the the number of scans; $a_m P_m(n)$ presents a term of the model and $e(n)$ indicates the additive noise. Using the orthogonal basis functions, Eq. (25) is equivalent to

$$y(n) = \sum_{m=1}^{M-1} g_m W_m(n) + e(n), \tag{27}$$

where $\overline{W_i(n)W_k(n)} = 0$ for $i \neq k$ and $\overline{W_i(n)W_k(n)} = \frac{1}{N}\sum_{n=1}^{N} W_i(n)W_k(n)$. g_m and $W_m(n)$ is demonstrated as

$$W_m(n) = P_m(n) - \sum_{r=1}^{m-1} \alpha_{mr} W_r(n), \quad m \geqslant 2, \tag{28}$$

and

$$\alpha_{mr} = \frac{\overline{P_m(n)W_r(n)}}{\overline{W_r^2(n)}}. \tag{29}$$

Here,

$$W_0(n) = P_0(n), \tag{30}$$

$$P_m(n) = \cos(\omega_r t_n), \tag{31}$$

$$P_{m+1}(n) = \sin(\omega_r t_n). \tag{32}$$

For a given orthogonal basis function $W_m(n)$, the coefficient estimate g_m which minimizes the mean square error (MSE) is given by

$$g_m = \frac{\overline{y(n)W_m(n)}}{\overline{W_m^2(n)}}, \tag{33}$$

and the MSE is

$$\overline{e^2} = \overline{\left[y(n) - \sum_{m=1}^{M-1} g_m W_m(n)\right]^2}. \tag{34}$$

When all the M basis functions have been selected, the coefficients a_m in Eq. (26) can be obtained

$$a_k = \sum_{m=k}^{M-1} g_m v_m, \tag{35}$$

where

$$v_k = 1, \quad v_m = -\sum_{r=k}^{m-1} \alpha_{mr} v_r,$$

and

$$m = k+1,\cdots,M-1. \tag{36}$$

Finally, the following model is

$$Y(n) = y(n)\beta(n) + e(n), \tag{37}$$

where $Y(n)$ is the observed time series, $e(n)$ shows the error term and β represents the parameter matrix. Then various statistic methods (e.g. t-test or F-test) can be performed to detect the brain activation region. Briefly, FOS is an iterative algorithm and can be used to detect the real-time activation signal.

1.2 Data-driven methods

Data-driven methods are complementary to model-driven methods. They can generate new hypotheses, separate and understand the nature of the mixture and then find non-trivial components of interest[18]. In this section, we introduce some popular data-driven methods such as principal component analysis (PCA), independent component analysis (ICA) as well as cluster analysis.

1.2.1 Principal component analysis PCA is a popular tool for the analysis of image data sets and is actively investigated in fMRI[19−22]. The objective is simplifying the data description by projecting the data vector onto the eigendirections in correspondence to the largest eigenvalues of the covariance matrix[22]. The image database is decomposed into several principal components by PCA. It describes the uncorrelated sequences in the image data set.

PCA is closely related to the singular value decomposition (SVD). For a data set $\boldsymbol{X}: N \times M, N > M$, its SVD is given by

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}, \tag{38}$$

where $\boldsymbol{U}$ is an $N \times M$ matrix and $\boldsymbol{V}$ is an $M \times M$

square matrix, both of them have orthogonal columns so that

$$\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U} = \boldsymbol{V}^{\mathrm{T}}\boldsymbol{V} = \boldsymbol{I}, \quad (39)$$

and $\boldsymbol{S}$ is an $M \times M$ diagonal matrix. The eigenvalues are the diagonal elements of $\boldsymbol{S}^2$, while $\boldsymbol{U}$ and $\boldsymbol{V}$ are the corresponding matrices of eigenvectors for $\boldsymbol{M}\boldsymbol{M}^{\mathrm{T}}$ and $\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}$, respectively. Postmultiplying $\boldsymbol{U}$ by $\boldsymbol{S}$ results in the matrix $\boldsymbol{P}$, which contains the principal component time series. PCA first measures the tendency of signals at all possible pairs of voxels, and then finds the orthogonal spatial patterns (eigenimages) capturing the greatest variance in the data. The first eigenimage represents the first component that can explain the data, the second eigenimage illustrates the second component and is orthogonal to the first eigenimage and so on. As illustrated in Fig. 1, there are some points scattered in the two-dimensional space. After PCA, we can find two directions that represent the first two components. The red line indicates the direction of the first principal component and the green is the second.

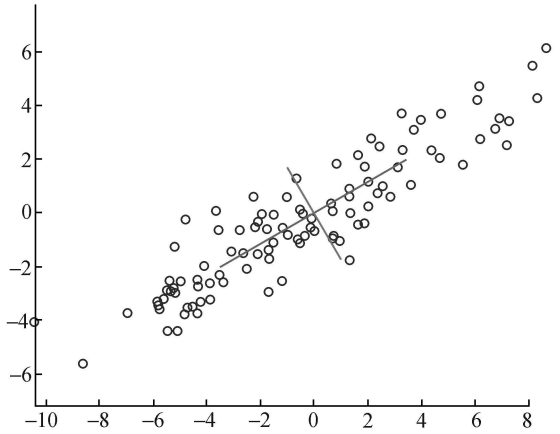

Fig. 1. Illustration of PCA. The red line indicates the direction of the first principal component and the green is the second.

For fMRI, the data set is arranged so that the temporal dimension is along the columns and the voxel-space dimension is along the rows (time by voxels). The most common use for PCA now is to reduce the dimensionality of the data while retaining the most information.

1.2.2 Independent component analysis In this section, we introduce the independent component analysis (ICA) and its applications in fMRI data analysis methods.

ICA is a signal processing technique. It separates a number of statistically independent sources that have been mixed linearly without further knowledge of their distributions or dynamics[19]. ICA was originally proposed for the blind source separation problem and used for solving the cocktail party problem.

In the cocktail party problem, m speakers are talking at the same time and recorded by N microphones. A given microphone is not placed to a specific speaker and is not shielded from other speakers. In this situation, each microphone records a mixture of all speeches. This is also called the blind source separation problem. The goal of ICA is to find a linear unmixing matrix $\boldsymbol{W}$ of the measurement matrix $\boldsymbol{X}$ (conversations recorded in microphones) which makes the rows of component matrix $\boldsymbol{Y}$ (the recovery of the sources) as independent as possible.

With the development of the neural network, a fast and efficient algorithm was developed as INFOMAX ICA[23]. It resolves the cocktail party problem by maximizing the information transfer in the unsupervised neural network. Bell et al.[23] proposed the original ICA algorithm to separate the super-Gaussian sources from the multi-channel measurements. Lee et al.[24] extended this method to model a more general source (super-Gaussian and sub-Gaussian). Both the original and extended ICA algorithms are based upon the assumption that the source distribution is sub-gaussian or super-gaussian, or their mixture. Xu et al.[25] proposed a non-linear approach to overcome this limitation so that there is no special demand on the source density when using ICA algorithm. Later, the scientists proposed other methods to select the model[26,27] for the fMRI data analysis. McKeown et al.[28] first adopted ICA algorithms for fMRI data analysis.

ICA in fMRI solves the same problem as it does in blind source separation. In fMRI study, ICA takes multi-channel measurements $\boldsymbol{X}$ as an $N \times M$ matrix, where N is the time points and M is the number of voxels. In a matrix form,

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S}, \quad (40)$$

where $\boldsymbol{X} = [x_1; x_2; \cdots; x_N]$ is the fMRI data set and $\boldsymbol{S} = [s_1; s_2; \cdots; s_N]$ is the source matrix of fMRI. According to Eq. (40), the j^{th} time course of $\boldsymbol{X}$ is represented by the linear combination of $a_i, 1 \leqslant i \leqslant N$, and their components are defined by the j^{th} column of $\boldsymbol{S}$, whose elements are statistically independent:

$$\begin{bmatrix} X_{1j} \\ X_{2j} \\ \square \\ X_{Nj} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \\ \square \\ A_{N1} \end{bmatrix} S_{1j} + \begin{bmatrix} A_{12} \\ A_{22} \\ \square \\ A_{N2} \end{bmatrix} S_{2j} + \cdots + \begin{bmatrix} A_{1N} \\ A_{2N} \\ \square \\ A_{NN} \end{bmatrix} S_{Nj}, \quad (41)$$

where $\boldsymbol{a}_i = [A_{i1}, A_{i2}, \cdots, A_{iN}]$, $\boldsymbol{a}_i$ is the i^{th} column of the mixing matrix $\boldsymbol{A}$. The BOLD signal $\boldsymbol{X}_{ij}$ is modelled as the summation of the individual independent component $[S_{1j}, S_{2j}, \cdots, S_{Nj}]$, which is weighed by the corresponding factors $[A_{i1}, A_{i2}, \cdots, A_{iN}]$.

Based on the mixing model, ICA directly solves the inverse problem which recovers source matrix $\boldsymbol{S}$ and finds the unmixing matrix $\boldsymbol{W}$ from the spatio-temporal fMRI data set $\boldsymbol{X}$. ICA iteratively determines the unknown unmixing matrix $\boldsymbol{W}$, which makes u_i as independent as possible. The unmixing principle is

$$\boldsymbol{U} = \boldsymbol{WX}, \quad (42)$$

where $\boldsymbol{U} = [u_1; u_2; \cdots; u_N]$. The process is illustrated in Fig. 2.

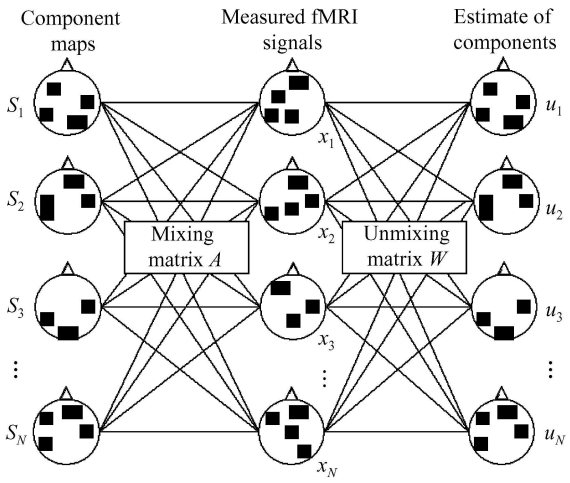

Fig. 2. Mixing and unmixing model in fMRI.

PCA and ICA are somehow related. PCA is a special case of ICA because the independence of random variables is more limited than the decorrelation. This is why PCA cannot separate the signals: they give uncorrelated components. In fMRI, ICA has been proved to be a better means for isolation and removal of structured noise, while PCA is superior for isolation and removal of random noise[19]. ICA can be applied to distinguish nontask-related signal components, movements and other artifacts. The consistently or transiently task-related fMRI can be activations defined as well. It is a highly promising method for fMRI data analysis and has been successfully employed by many researches. PCA is an important preprocessing step before ICA execution.

Finally, we make a brief introduction of the INFOMAX ICA in fMRI[28], which is the most often-used criterion for ICA. First, initialize $\boldsymbol{W}$ to the identity matrix $\boldsymbol{I}$ and denote $\boldsymbol{X}_s = \boldsymbol{PX}$, where $\boldsymbol{P} = 2\langle \boldsymbol{XX}^{\mathrm{T}} \rangle^{-\frac{1}{2}}$, and $\langle \boldsymbol{XX}^{\mathrm{T}} \rangle$ is the covariance matrix of the fMRI data matrix $\boldsymbol{X}$. Then compute

$$\Delta \boldsymbol{W} = [\boldsymbol{I} + (1 - 2y_i)\boldsymbol{U}^{\mathrm{T}}]\boldsymbol{W}, \quad (43)$$

where $\boldsymbol{U} = \boldsymbol{WX}_s$ and $y_i = g(u_i) = \dfrac{1}{1 + \mathrm{e}^{-u_i}}$. Using Eq. (43) to update $\boldsymbol{W}$, $\boldsymbol{W} \leftarrow \boldsymbol{W} + \Delta \boldsymbol{W}$ until the stop criterion is approached (e. g. root mean square changes for all elements $< 10^{-6}$).

1.2.3 Cluster analysis Another type of data-driven approach is cluster analysis, which groups image pixels based on the similarities defined by the chosen distance measurement. The result consists of a partition of the data (known as clusters) and the corresponding cluster center. Each cluster contains voxels with the similar signal time courses and the set of cluster centers represents the structure in the data. Most commonly used cluster methods in fMRI are c-means[29—32], fuzzy cluster analysis (FCA)[33—36], dynamical cluster analysis (DCA)[37], and deterministic annealing[38]. One of the most important issues in the cluster analysis is to estimate the number of the clusters as Charles et al.[39] proposed. They first evaluate the signal subspace, and the maximum likelihood estimation is implemented via the expectation-maximization algorithm under several assumptions which form the problem into the Gaussian mixture model. Then a minimum description length (MDL) criterion is proposed to figure the number of clusters. And this method performs better than the traditional manner such as c-means and FCA.

Self-organizing map (SOM) is also one of the topology-preserving clustering methods. It classifies the time signal of the brain into several patterns according to the temporal similarity of the signals. SOM is understood as a neighborhood-preserving mapping of highly dimensional data onto a two-dimensional lattice while preserving the topological

structure of a data set. This map represents the cluster centers which are updated by taking randomly selected features of the data set. The difference between the classic clustering methods and SOM is that the incorporation of a new feature into the nearest cluster has an impact on the neighboring.

Many researchers adopt SOM to analyze fMRI data set[40—43]. In SOM technique, the number of nodes used is predetermined by the user under a few general guidelines. For example, if the number of the nodes is too small, the flexibility of this manifold and the goodness of the fit are poor. While if the number of nodes is too large, more computational demanding is needed. William et al.[43] proposed a method to solve this problem. They use a cluster merging technique based on examining the reproducibility of the fMRI data across epochs to merge certain SOM nodes (called "candidate nodes") and the resulting "super nodes" give the time course template of the potential interest which is used in the subsequent traditional template-based analysis. Here, we briefly discuss the basic SOM algorithm in fMRI[42]. In pseudocode, SOM clustering works as follows:

1) fix SOM dimension along with training parameters;

2) initialize each node;

3) randomly select a pixel time course (PTC);

4) find the cluster center with minimal distance to the PTC;

5) move this cluster center towards the selected PTC;

6) elastically move the neighboring cluster center on the SOM;

7) if the stopping criteria are not yet satisfied, go to 3).

When map training is processing, the amount of the cluster center is reduced and the neighborhood of the current center is reduced as well. This behavior is modeled by a neighborhood function h_{ci}, when randomly selecting a PTC x with its nearest cluster center c_x, the update function for all cluster centers c_i takes the form:

$$c_i(t+1) = c_i(t) + h_{ci}(t)(x(t) - c_i(t)), \tag{44}$$

where $h_{ci}(t)$ should be

$$h_{ci}(t) = h(d(r_c - r_i), t), \tag{45}$$

and $r_c \in \mathcal{R}^2$, $r_i \in \mathcal{R}^2$ are the location vectors of SOM nodes c_x and c_i in the array, and $d(\cdot)$ is the Euclidean distance.

1.2.4 Clustered cNMF Clustered cNMF stands for clustered constrained non-negative matrix factorization. Non-negative matrix factorization (NMF) is a relatively new technique proposed for dimensionality reduction[44,45]. It employs the Poisson statistics as a noise model and preserves a lot of structure of the original data. NMF is motivated mainly because the fMRI data is positively defined and NMF is based on positive restrictions, meaning that NMF can be a suitable method for such a problem. In addition, NMF computation is based upon the simple iteration algorithm, and it provides a nice simple learning rule, which is guaranteed to converge monotonically. Recently, Wang et al.[46] proposed a new method as clustered constrained NMF to estimate the statistically distinct neural responses in a sequence of fMRI. They utilized an improved objective function which is more suitable for the task-related brain activation detection and placed particular emphasis on the initialization of the cNMF algorithm. Finally, they use the K-means algorithm, and the information theoretic criterion of minimum description length (MDL) is used to estimate the number of clusters. The algorithm can be summarized by the following steps:

1) initialize K with a large number $K_{\max}$;

2) group the data set using the k-means algorithms;

3) calculate the MDL(K);

4) if $K < K_{\min}$, go to step 5); otherwise, $K = K - 1$, then go to step 2);

5) select K and its corresponding parameters that results in the smallest value of the MDL criterion;

6) execute the cNMF algorithm.

The authors developed a new method to fMRI data analysis based on NMF, and therefore provided a new framework of the application of cNMF to the task-related fMRI data analysis.

2 Applications in fMRI

fMRI, as a window into the brain, has been successfully applied in many research areas, such as maps of somatosensory systems[47], higher cognitive processes[48,49] and human memory systems[50–52]. Recently, scientists in China have done a lot of work to study the human emotion[53,54] and the acupuncture[55] which attracts much attention.

Yang et al.[53] discussed the gender differences during human emotion processing using International Affective Picture System (IAPS) in an event-related fMRI study. The results indicate that the activation of left postcentral gyrus and left inferior parietal lobule by positive pictures compared with neutral pictures was observed in male but not in female when $p < 0.005$. Whereas, female showed a significant BOLD signal response in bilateral putamen, right amygdala, bilateral para-hippocampal gyrus, bilateral occipital gyrus (BA19/18), right temporal gyrus and the bilateral cerebellum (Fig. 3(a)). In the group analysis of the brain activation (when comparing positive pictures versus neutral pictures in the contrast male > female, $p < 0.005$), left middle occipital gyrus, left middle frontal gyrus (BA10) and right para-hippocampal gyrus remained significant. Negative emotional pictures elicited a significant activation of left thalamus, bilateral pons, left midbrain, left putamen, left cerebellar tonsil and the right superior temporal gyrus (BA22) in female when the threshold was set to $p < 0.001$. There is no significant activation found in male. Based on a hypothesis of amygdala and insula activation to aversive stimuli, brain activation test was conducted at the significance level $p < 0.01$. The activation was found only in female in left amygdala and right insula as shown in Fig. 3(b).

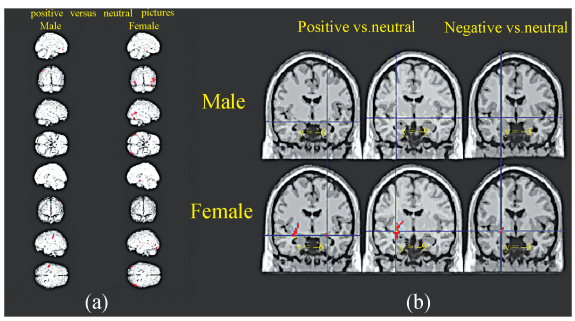

Fig. 3. Brain activation in male and female when viewing positive versus neutral pictures and negative versus neutral pictures.

This work has revealed that female has more activation by negative stimuli compared to neutral in midline limbic structures, including thalamus, midbrain, and cerebellum. One possible reason is that female showed more attention to the feeling state engendered by emotional stimuli and greater overt response, possibly for social reasons. They tested brain activation at the significance threshold of $p < 0.01$ in order to find whether amygdala and insula are activated. Only female demonstrated left amygdala activation and right insula activation. This study indicates that female has more sensitive central processing to aversive materials in general. This result may explain the higher depression rate in women. In general, female shows more frequent activation in the basal ganglia under either positive or negative stimuli. In positive emotions, basal ganglia may play a pivotal role in broadening the repertoire of accessible thoughts and actions which result in exploratory behavior and skill-building, leading to the activation of a number of functional loops in the basal ganglia that implement a wide range of thoughts and behavior. In aversive situation, the basal ganglia may develop a specific role for threat, leading to more focal stereotyped responses.

This is one of the few studies using the images of IAPS to illustrate gender differences in recognition of positive and negative emotions. Consistent with previous findings, it concludes that male and female employ different sets of neural correlations to process positive and negative pictures. It is more noticeable when they were processing negative pictures than positive pictures.

3 Conclusions and future studies

fMRI is a noninvasive examination with the excellent spatial resolution, more and more scientists choose it as the research means. And many algorithms have been performed to reveal the intrinsic relationship between the neuronal activity and the measured signals. Most of the above mentioned methods have dealt with linear systems. This is obviously inaccurate because of fMRI principles. We plan to find the intrinsic geometry of the signal space which take an appropriate measurement. Another approach is the threshold selection which is the key point in fMRI signal detection. Different threshold leads to different results such as the number of the activated voxels or areas. A fine algorithm must contain a fine threshold selection criterion. The third question is the smoothing in space and time. Smoothing in space enhances the SNR of the data and allows intersubject averaging

by blurring differences in gyral anatomy between subjects. On the other hand, fMRI has high anatomical resolution and so there is a tradeoff between the degree of smoothing and the spatial resolution. An appropriate noise reduction technique should be further considered. The functional integration is another demand that we should seriously deliberate. Brain works by function segregation and integration. Function segregation assumes that different brain area works individually. This is suitable for simple tasks such as in Broca's language areas. While brain works, many areas are involved in the task and how these areas connected and what is the relationship between them should be investigated. Currently, most researches have focused on the functional segregation instead of integration. More efforts should be taken for functional integration. Finally, compared to the electroencephalogram (EEG) and magnetoencephalography (MEG), the temporal resolution is rather poor in fMRI. How to combine EEG/MEG and fMRI techniques to improve both the spatial and temporal resolution is a challenging task[56—60].

Acknowledgement The authors would like to thank Prof. Liu Ying for the proofreading and the helpful suggestions.

References

1 Roy C.S. and Sherrington C.S. On the regulation of blood supply of the brain. Journal of Physiology, 1890, 1:85—108.
2 Ogawa T., Tank D.W., Menon R. et al. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. In: Proceedings of the National Academy of Sciences, 1992, 89:5951—5955.
3 Belliveau J.W., Kennedy D.N., McKinstry R.C. et al. Functional mapping of the human visual cortex by magnetic resonance imaging. Science, 1991, 254:716—719.
4 Bandettini P. A., Wong E. C., Tikofsky R. S. et al. Time course EPI of human brain function during task activation. Magnetic Resonance in Medicine, 1992, 25(2):390—397.
5 Glover G. H. Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage, 1999, 9:416—429.
6 Ward B. D. Deconvolution analysis of fMRI time series data, http://brainimaging.waisman.wisc.edu/, June, 2001.
7 Ashburner J. and Friston K.J. Nonlinear spatial normalization using basis functions. Human Brain Mapping, 1999, 7(4):254—266.
8 Petersson K.M., Nichols T.E., Poline J.B. et al. Statistical limitations in functional neuroimaging Ⅱ. Signal detection and statistical inference. Phil. Trans. R. Soc. Lond., 1999, 354(1387):1261—1281.
9 Kiebel S. J., Goebel R. and Fristonk K.J. Anatomically informed basis functions. NeuroImage, 2000, 11(1):656—667.
10 Frison K.J., Jezzard P. and Turner R. Analysis of functional MRI time series. Human Brain Mapping, 1994, 1(2):153—171.
11 Friston K.J., Frith C.D., Turner R. et al. Characterizing evoked hemodynamics with fMRI. NeuroImage, 1995, 2(2):157—165.
12 Friston K.J., Holmes A.P., Poline J.B. et al. Analysis of fMRI time series revisited. NeuroImage, 1995, 2(1):45—53.
13 Friston K.J., Holmes A.P. and Worsley K.J. Statistical parametric maps in functional imaging: a general linear approach. Human Brain Mapping, 1995, 2(1):189—210.
14 Buchel C., Wise R.J. and Mummery C.J. Nonlinear regression in parametric activation studies. NeuroImage, 1996, 4(1):60—66.
15 Friston K.J., Josephs O., Rees G. et al. Nonlinear event-related responses in fMRI. Magnetic Resonance in Medicine, 1998, 39(1):41—52.
16 Kellman P., Gelderen P.V., Zwart J.A. et al. Method for functional MRI mapping of nonlinear response. NeuroImage, 2003, 19(1):190—199.
17 Li X.F., Tian J., Wang X.X. et al. Fastorthogonal search for modelling nonlinear hemodynamic response in fMRI. In: SPIE International Symposium Medical Imaging, San Diego, California, USA, February 11—17, 2004, 219—226.
18 Martin J. M., Lars K. H. and Terrence J. S. Independent component analysis of functional MRI: what is signal and what is noise? Current Opinion in NeuroBiology, 2003, 13(5):620—629.
19 Christopher G. T., Richard A. H. and Ravi S. M. Noise reduction in BOLD-based fMRI using component analysis. NeuroImage, 2002, 17(3):1521—1537.
20 Andersen A.H., Gash D.M. and Avison M.J. Principal component analysis of the dynamic response measured by fMRI: a generalized linear systems framework. Magnetic Resonance Imaging, 1999, 17(6):795—815.
21 Friston K.J., Philips J., Chawla D. et al. Revealing interactions among brain systems with nonlinear PCA. Human Brain Mapping, 1999, 8(2):92—97.
22 Hansen L.K., Larsen J., Nielsen F. et al. Generalizable patterns in neuroimaging: how many principla components? NeuroImage, 1999, 9(1):534—544.
23 Bell A. J. and Sejnowski T. J. An information-maximization approach to blind source separation and blind deconvolution. Neural Computation, 1995, 7(6):1129—1159.
24 Lee T., Girolami A.M. and Seinowski T.J. Independent component analysis using an extended informax algorithm for mixed subgaussian and super-gaussian sources. Neural Computation, 1999, 11(2):409—433.
25 Xu L., Cheung C.C. and Amari S. Learned parametric mixture based on ICA algorithm. Neurocomputing, 1998, 22(1—3):69—80.
26 Beckmann C.F., Noble J. and Smith S.M. Investigating the intrinsic dimensionality of FMRI data for ICA. NeuroImage, 2001, 13:S76.
27 Calhoun V.D., Adali T., Pearlson G.D. et al. A method for making group inferences from functional MRI data using independent component analysis. Human Brain Mapping, 2001, 14(3):140—151.
28 McKeown M. J., Makeig S., Brown G.G. et al. Analysis of fMRI data by blind separation into independent spatial components. Human Brain Mapping, 1998, 6(3):160—188.
29 Wong E.C., Buxton R.B., Blaettler D. et al. Unbiased identification of funtionally disinct groups of pixels in functional MRI using cluster analysis in Fourier space. In: The Third Annual Meeting of the International Society for Magnetic Resonance in Medicine. Rice France, August 19—25, 1993.
30 Ding X., Tkach J., Ruggieri P. et al. Analysis of time-course functional MRI data with clustering method without the use of reference signal. In: The Second Annual Meeting of the International Society for Magnetic Resonance in Medicine. San Francisco, 1994.

31 Goutte C., Toft P., Rostrup E. et al. On clustering fMRI time series. NeuroImage, 1999, 9(3): 298—310.

32 Balslev D., Nielsen F. A., Frutiger S. A. et al. Cluster analysis of activity-time series in motor learning. Human Brain Mapping, 2002, 15(3): 135—145.

33 Golay X., Kollias S., Meier D. et al. Optimization of a fuzzy clustering technique and comparison with conventional post processing methods in fMRI. In: The Fourth Annual Meeting of the International Society for Magnetic Resonance in Medicine. New York, April 27—May 3, 1996.

34 Golary X., Kollias S., Stoll G. et al. A new correlation-based fuzzy logic clustering algorithm for fMRI. Magnetic Resonance in Medicine, 1998, 40(2): 249—260.

35 Baumgartner R., Windischberger C. and Moser E. Quantification in functional magnetic resonance imaging: fuzzy clustering vs. correlation analysis. Magnetic Resonance Imaging, 1998, 16(2): 115—125.

36 Fadili M. J., Ruan S., Bloyet D. et al. A multistep unsupervised fuzzy clustering analysis of fMRI time series. Human Brain Mapping, 2000, 10(4): 160—178.

37 Baune A., Sommer F. T., Michael E. et al. Dynamical cluster analysis of cortical fMRI activation. NeuroImage, 1999, 9(5): 477—489.

38 Wismuller A., Lange O., Dersch D. R. et al. Cluster analysis of biomedical image time-series. International Journal of Computer Vision, 2002, 46(2): 103—128.

39 Chen S., Bouman C. A. and Lowe M. J. Clustered components analysis for functional MRI. IEEE Transactions on Medical Imaging, 2004, 23(1): 85—98.

40 Chuang K. H., Chiu M. J., Lin C. C. et al. Model-free functional MRI analysis using Kohonen clustering neural network and fuzzy c-means. IEEE Transactions on Medical Imaging, 1999, 18(12): 1117—1128.

41 Ngan S. C. and Hu X. P. Analysis of functional magnetic resonance imaging data using self-organizing mapping with spatial connectivity. Magnetic Resonance in Medicine, 1999, 41(5): 939—946.

42 Fischer H. and Hennig J. Neural network-based analysis of MR time series. Magnetic Resonance in Medicine, 1999, 411: 124—131.

43 Ngan S. C., Yacoub E. S., Auffermann W. F. et al. Node merging in Kohonen's self-organizing mapping of fMRIdata. Intelligence in Medicine, 2002, 25(1): 19—33.

44 Lee D. D. and Seung H. S. Leanring the parts of objects by non-negative matirx factorization. Nature, 1999, 401: 788—791.

45 Lee D. D. and Seung H. S. Algorithms for non-negative matrix factorization. Advances in Neural Information Processing Systems, 2000, 13: 556—562.

46 Wang X. X., Tian J., Li X. F. et al. Detection the brain activations by constrained non-negative matrix factorization from task-related fMRI. In: Proceedings of SPIE Medical Imaging, San Diego, California, USA, February 11—71, 2004, 675—682.

47 Reed C. L., Klatzky R. L. and Halgren E. What vs. where in touch: an fMRI study. NeuroImage, 2005, 25(1): 718—726.

48 Remijnse P. L., Nielen M. A., Uylings B. M. et al. Neural correlates of a reversal learning task with an affectively neutral baseline: an event-related fMRI study. NeuroImage, 2005, 26(1): 609—618.

49 Koelsch S., Fritz T., Schulze K. et al. Adults and children processing music: an fMRI study. NeuroImage, 2005, 25(1): 1068—1076.

50 Maguire E. A., Frith C. D., Rudge P. et al. The effect of adult-acquired hippocampal damage on memory retrieval: an fMRI study. NeuroImage, 2005, 27(1): 146—152.

51 Kensinger E. A. and Schacter D. L. Retrieving accurate and distorted memories: neuroimaging evidence for effects of emotion. NeuroImage, 2005, 27(1): 167—177.

52 Sarah E. D., Carter W., Eveline A. C. et al. Retrieving rules for behavior from long-term memory. NeuroImage, 2005, 26(7): 1140—1149.

53 Yang L., Tian J., and Wang X. X. Gender differences in the processing of standard emotional visual stimuli: an event-related fMRI study. In: Proceedings of the 90th Scientific Assembly and Annual Meeting of Radiological Society of North America (RSNA 2004), Chicago, USA, Nov. 28—Dec. 3, 2004 (oral report).

54 Wang X. X., Tian J., Yang L. et al. Functional connectivity in amygdala for affective pictures perception studied with functional magnetic resonance imaging. In: Proceedings of 10th Annual Meeting of the Organization for Human Brain Mapping, MIAMI, Florida, USA, June 13—17, 2005.

55 Qin W., Tian J. and He H. G. Brain activation by moxibustion: an fMRI study. In: Proceedings of 10th Annual Meeting of the Organization for Human Brain Mapping, MIAMI, Florida, USA, June 13—17, 2005.

56 Dale A. M. and Sereno M. I. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. Journal Cognitive Neuroscience, 1993, 5(2): 162—176.

57 Liu K., Belliveau J. W. and Dale A. M. Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. In: Proceedings of the National Academy of Sciences USA, 1998, 8945—8950.

58 Dale A. M., Liu A. and Fischl B. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron, 2000, 26(1): 55—67.

59 Babiloni F., Babiloni C., Carducci F. et al. Multimodal integration of high-resolution EEG and functional magnetic resonance imaging data: a simulation study. NeuroImage, 2003, 19(1): 1—15.

60 Yang L., Tian J., Wang X. X. et al. Gender differences in the processing of standard emotional visual stimuli: integrating ERP and fMRI results. In: SPIE International Symposium on Medical Imaging, San Diego, California, USA, February 12—17, 2005, 648—656.