Correspondence

Delay Correlation Subspace Decomposition Algorithm and Its Application in fMRI

Huafu Chen, Dezhong Yao*, Wufan Chen, and Lin Chen

***Abstract*—This paper reports a new delay subspace decomposition (DSD) algorithm. Instead of using the canonical zero-delay correlation matrix, the new DSD algorithm introduces a delay into the correlation matrix of the subspace decomposition to suppress noises in the data. The algorithm is applied to functional magnetic resonance imaging (fMRI) to detect the regions of focal activities in the brain. The efficiency is evaluated by comparing with independent component analysis and principal component analysis method of fMRI.**

***Index Terms*—Functional magnetic resonance imaging (fMRI), independent component analysis (ICA), noise suppression, principal component analysis (PCA), subspace decomposition.**

I. Introduction

Functional magnetic resonance imaging (fMRI) has emerged as a useful tool in the study of brain function. Much of the work in fMRI data analysis has revolved around the creation of statistical maps and the detection of activation at different voxels of brain images [1]–[3]. Some researchers have attempted to address the issue by using parametric methods which usually assume specific signal shapes (Poisson, Gaussian, Gamma, etc.) and attempt to extract the associated parameters for which the data best fit [4], [5]. Others have taken a linear systems approach in which the response is modeled as an impulse response convolved with the stimulus reference function [6], [7]. Such as the well-known general linear model (statistical parameter mapping, SPM) method which is based on known experiment pattern information [8], [9]. These methods are classified as model-driven method because some prior information such as the response function of a stimulus was utilized in the algorithm. Recently, exploratory data-driven methods got more and more attention, as they do not require any knowledge and hypothesis about the paradigm or the hemodynamic response function. The representative algorithm of the data-driven methods is the independent component analysis (ICA) method [10], [11]. ICA can be used to get independent components from mixed signals, with the constraint that the separated components are mutually independent. As a high-order statistics based method, ICA can suppress white and colored Gaussian noises automatically [12]. The basic idea of subspace decomposition named multiple signal classification (MUSIC) [13], published in 1986 on the radar signal detection problem, is to decompose the data space of the autocorrelation matrix of an array recordings, by such an algorithm as the singular value decomposition (SVD), into signal subspace and noise subspace, then conducted are the signal projection distribution in one of the two subspaces to get the location information of the desired sources. This method was firstly introduced to brain function detection in 1992 as a magnetoencephalography inverse algorithm [14].

The first application of MUSIC in fMRI was reported in 1996, and was designed to detect cortical activities from fMRI time-course data with a forward and backward covariance averaging [15]. Different from our current work, this method was a MR imaging process involved algorithm and was only evaluated by simulation data. Our method reported here is an image processing method totally based on the obtained fMRI data; thus, it may be sorted to the class of data-driven fMRI algorithms.

This paper is organized as follows: The proposed method for delay subspace decomposition (DSD) is introduced in the second section. Simulation data and fMRI experiment are then introduced in the third section. The result is presented in the forth section and finally followed by a discussion.

II. DSD Method

We assume that the neural activities that produce the fMRI are all focal activities, and fMRI data are composed of functional active signals and noisy background activities; thus, the MRI measured data at voxel p $y_p(t)$ is a sum of the functional activities $s_p(t)$ and noises $\mathrm{e_p(t)}$

$$y_p(t) = s_p(t) + e_p(t), \quad t = 1, \ldots, N, \quad p = 1, \ldots, P \tag{1}$$

where N is the number of the temporal sample points, P is the number of voxels. In a matrix notation, we have

$$y = s + e \tag{2}$$

where $y \in R^{P\times N}$, $s \in R^{P\times N}$ and $e \in R^{P\times N}$.

The delay correlation matrix of the measured fMRI data y is defined as

$$R(\beta) = y(t)y^T(t+\beta) \tag{3}$$

and

$$\begin{aligned} R(\beta) &= y(t)y^T(t+\beta) \\ &= s(t)s^T(t+\beta) + s(t)e^T(t+\beta) \\ &\quad + s(t+\beta)e^T(t) + e(t)e^T(t+\beta) \end{aligned} \tag{4}$$

where T denotes matrix transpose; β is the delay time. When $\beta = 0$, this matrix reduces to the normal case utilized by the canonical MUSIC algorithm [13]. As the source and evolution process of a stimulus induced signal and the background noises are intrinsic different, we assume their mutual correlation, $s(t)e^T(t+\beta) + s(t+\beta)e^T(t)$, may be neglected when compared to the self-correlation of signal or noise, then (4) changes to

$$R(\beta) \approx s(t)s^T(t+\beta) + e(t)e^T(t+\beta). \tag{5}$$

Furthermore, similar to the smoothness constraint imposed by the differential operator in the well-known Tikhonov regularization [16],

Manuscript received May 3, 2005; revised August 30, 2005. This work was supported in part by the 973 Project under Grant 2003CB716106, in part by the NSFC under Grant 90208003 and Grant 30200059, in part by the Doctor training Fund of MOE, PRC (TRAPOYT), in part by the Fok Ying Tong Education Foundation under Grant 91041, and in part by the University of Electronic Science and Technology of China (UESTC) through the Middle-youth academic leader training plan. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was P. Thompson. *Asterisk indicates corresponding author.*

H. Chen and L. Chen are with the School of Life Science & Technology, School of Applied Math, University of Electronic Science and Technology of China, Chengdu 610054, China. They are also with The Key Laboratory of Cognitive Science, Graduate School and Institute of Biophysics, Chinese Academy of Science 100101, China.

*D. Yao is with the School of Life Science & Technology, School of Applied Math, University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: Dyao@uestc.edu.cn).

W. Chen is with the School of Biomedical Engineer, Southern Medical University, Guangzhou 510515, China.

Digital Object Identifier 10.1109/TMI.2005.859211

we assume the meaningful stimulus induced response is much more smooth than various randomly generated noises, thus by introducing the delay parameter β, we may have the signal-to-noise-ratio (SNR) of the delay correlation matrix $R(\beta)$ being much larger than SNR of the original correlation matrix $R(0)$. How much SNR enhances depends on the intrinsic smooth difference and the delay parameter β, and for some extreme cases, we may have a $R(\beta)$ where the noises are almost removed completely as shown in the following example. In general, we are not to find a delay parameter to remove the noise completely, but a small β that enhance SNR for the following subspace decomposition.

Conducting a SVD on R

$$R = U\Sigma V^T \tag{6}$$

where $\Sigma = \mathrm{diag}[\lambda_1, \ldots, \lambda_P]$ is the eigenvalue of matrix R or singular values of the matrix, and $\lambda_{\mathrm{i}} > \sigma^2, \mathrm{i} = 1, \ldots, \mathrm{L}; \lambda_{\mathrm{i}} \approx \sigma^2, \mathrm{i} = \mathrm{L}+1, \ldots, \mathrm{P}; \sigma^2$ is the remaining noise variance which may a very small value close to zero. The matrixes U and V are composed of the corresponding eigenvectors. According to the singular value distribution, the matrix U or V can be decomposized into signal subspace and noise subspace, such as $U = [S|G] = [u_1, \ldots, u_L | u_{L+1}, \ldots, u_P]$, where $S = [u_1, \ldots, u_L]$ and $G = [u_{L+1}, \ldots, u_P]$ corresponds to the signal subspace and the noise subspace, respectively.

The projection of the orginal signal dataset to the signal subspaces defined as

$$\bar{S} = S^T Y. \tag{7}$$

As the signal subspace S is composed of the eigenvectors corresponding to the first L eigenvalues, $\bar{S}$ contains the principal signals of the original data. For each voxel singal y_p, we define a measure as

$$f(p) = \frac{\|\bar{S} y_p\|}{\|\bar{S}\| \|y_p\|}. \tag{8}$$

Equation (8) means a focal maximum will appear if the signal y_p is similar to the principal signals, or say the signal y_p is located in the signal subspace. Thus the space distribution of the measure $f(p)$ will give the information of source locations. Compared with the principal component analysis (PCA) algorithm of fMRI, it is clear that the PCA algorithm is just the special case with $L = 1$ and $\beta = 0$ [17].

Based on the above equations, the procedure of the DSD algorithm is as follows:

1) determining a proper delay paramter according to the intrinsic difference of the signals and noises;
2) generating the signal subspace;
3) calculating the signal subspace projection by (7);
4) calculating the imaging measure $f(p)$ and localizing the maxima from the distribution of $f(p)$.

The key step of DSD is the choice of L. For $\beta = 0$, in the case of the canonical subspace decompositon [14], we have $\lambda_{\mathrm{i}} = \delta^2, \mathrm{i} = \mathrm{L}+1, \ldots, \mathrm{P}$, and a big $\sigma^2 (\beta = 0)$ may mask the truning point on the singular value distribution, thus making it very difficult to choose a correct L. By using the delay correlation, a nonzero β may induce an $\mathrm{ee}^{\mathrm{T}}(\beta)$ much smaller than $\mathrm{ee}^{\mathrm{T}}(\beta = 0)$ especially when the random noises are of wide spectrum. Meanwhile, as many meaningful signals such as a stimulus induced fMRI signal $s(t)$ is usually a smoothly changed signal of narrow spectrum, $s(t)s^T(t+\beta)$ may still keep the most information when β is not too large. In our work, the idea behind specifying β is to reduce noise significantly without decreasing the signal correlation too much. In particular, based on simulations, we choose a delay parameter β which corresponds to the biggest signal delay autocorrelation (DA) and noise DA ratio and the related signal DA still is larger than 0.6 times the signal zero DA.

For comparison, spatial ICA was utilized in this work. Details of spatial ICA can be found in [10] and [18].

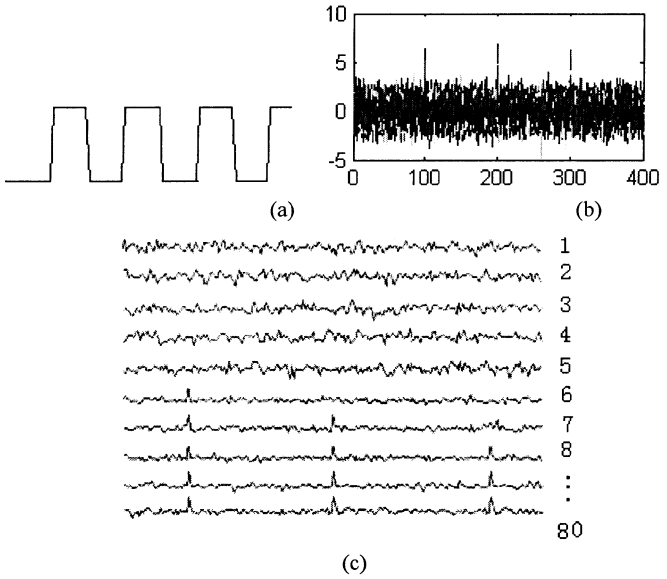

Fig. 1. Simulation data. (a) Block signal simulating a fMRI stimulus pattern; (b) An fMRI dataset with three activation locations; the horizontal-axis is the order number of voxel, the vertical axis is the intensity value of the signal. (c) The simulated fMRI dataset with three asynchronous activations in Gaussian noisy background with different delay-times.

III. Simulation Study

A. Stimulated Data

The *in vivo* fMRI dataset is generally composed of brain functional activations and background noises which are assumed to be Gaussian noise in our simulation.

Suppose that the section under investigation is composed of 400 voxels (20×20 voxels fMRI image) and a dataset composed of 400 stochastically generated samples by the Gaussian noise generator in Matlab is taken as the spontaneous background signals of this section at one sampling time point. The generation procedure was repeated for 80 times. Finally, the simulated dataset has a dimension of 400×80. A block signal shown in Fig. 1(a) is added to the generated background noise on voxel 100 with different SNR defined as $\mathrm{SNR} = 10 * \log(\mathrm{std}(\mathrm{signal})/\mathrm{std}(\mathrm{noise}))$ dB, and other two block signals with one and two sample point lagged respectively to the first one are added on voxels 200 and 300, respectively.

In summary, we get a 400-voxel fMRI dataset, in which there are three asynchronous excitation points with different SNR. The three activation locations are shown in Fig. 1(b) where $\mathrm{SNR} = 1.5$ dB. Fig. 1(c) shows the 80 time points of the 400 voxels in Fig. 1(a). The first to the fifth channel in Fig. 1(b) denote the resting state. The sixth channel denotes that there is an activation point, the seventh channel has two activation points, the eighth channel has three activation points. The three activation points of fMRI stimulation dataset are asynchronous because the first one begins to excite at the sixth time point, the second begins at the seventh point and the third at the eighth point.

B. Result

1) Result of Spatial ICA: The dataset was decomposed into 80 signals by the spatial ICA method as shown in Fig. 2. Where the excited points can be respectively found in the first, second and third components, their corresponding images are the fMRI results of the three assumed asynchronous activations, and other components in Fig. 2 are the other fMRI noises. However, if we do not know the number of active regions, the choice of components will be a challenge problem [12],

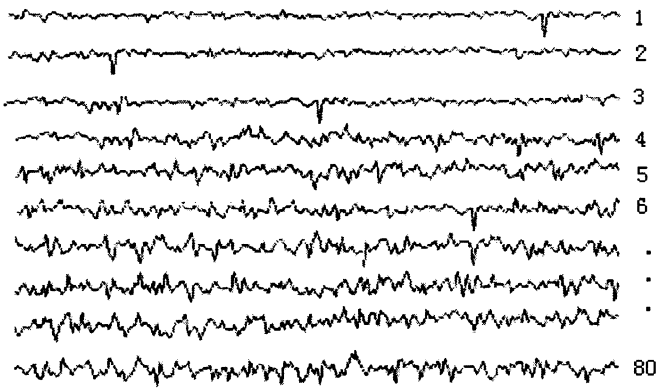

Fig. 2. Spatial ICA result.

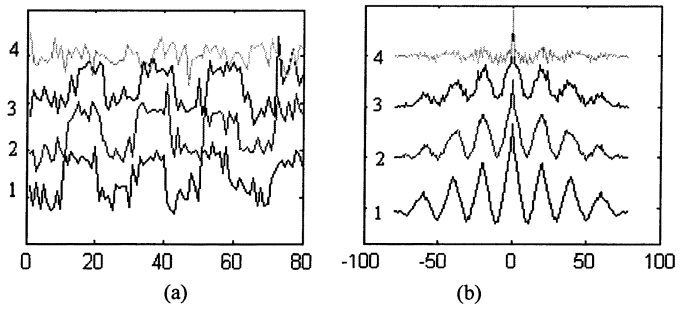

Fig. 3. Noisy signals and DA. (a) Noisy signals, the horizontal axis is the order number of time sample point. (b) DA, the horizontal-axis is the number of delay time point.

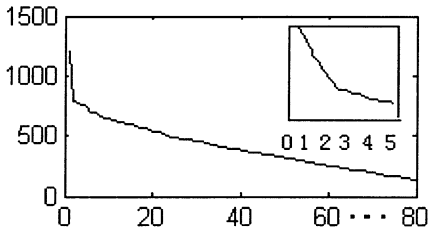

Fig. 4. Cure of the singular values. The horizontal-axis is the order number of the singular values, the vertical axis is the value.

for example, the sixth component in Fig. 2 may be wrongly chosen as a meaningful signal.

2) Result of DSD: There are three main steps in DSD: 1) determining the delay parameter β; 2) choosing the L which determines the decomposition of the signal space and noise space; 3) localizing the activation points.

The first to third signal in Fig. 3(a) are the signals of the three active voxels in Fig. 1(b) and the fourth is noise, their DAs are shown in Fig. 3(b). The normalized DA values of the fourth noise temporal process are 1.00, 0.197, 0.101, and 0.029 when parameter β is 0, 1, 2, and 3, respectively. Meanwhile, the normalized DA values of the first signal temporal process are 1.00, 0.892, 0.794, and 0.696 when parameter β is 0, 1, 2, and and 3, respectively. As noted above, we choose a delay parameter β based on the biggest signal DA and noise DA ratio when signal DA still is larger than 0.6. Here, the ratios are $(1.0)/(1.0) = 1.0, (0.892)/(0.197) = 4.53, (0.794/0.101) = 7.9, (0.696)/(0.029) = 23.0$; thus, we choose $\beta = 3$. Fig. 4 shows the distribution curve of the singular values with an amplified subfigure for the local information which indicates $\mathrm{L} = 3$. Fig. 5 (left column) shows the simulated fMRI data with $\mathrm{SNR} = 1.5$ dB [Fig. 5(top)], 1.0 dB [Fig. 5 (middle), and 0.7 dB [Fig. 5 (bottom)]. Fig. 5 (middle column) shows the corresponding localization results of DSD according to (7) with $\beta = 3$ and $L = 3$. Fig. 5 (right column) shows the results by DSD with $\beta = 0$ and $L = 1$, thus the result of PCA, correspondingly. The three spikes whose intensity is larger than 20 is actually the assumed three activation points. This figure demonstrates that the proposed method localizes the activation points better than PCA, and this becomes more evident as the noise level increases

For comparison, the feasabilities of DSD and spatial ICA for dataset with different SNR are evaluated by simulated data (the details of spatial ICA method are in [18]). Table I shows the results. We find that the result of spatial ICA and DSD are almost the same in synchronous brain activation. While DSD is obviously better than spatial ICA in asynchronous situation. DSD is validity when SNR is larger than 0.3 dB for synchronous activations or 0.6 dB for asynchronous activations, while the spatial ICA suffers from the choice problem of separated components, the required SNR is higher than DSD did.

For inspecting the feasibilities of DSD for different noises, a fMRI data of a rest state was utilized as a non-Gausssian noise instead of the Gaussian noise in above simulation test, the results (omitted) show the DSD method is valid as the Gaussian noise case.

IV. Real fMRI Data Test

A. Human Data Acquisition

The fMRI data was collected by Beijing Cognitive Lab. in Beijing hospital. The stimulus was a red illuminant point presented at the center of the visual field with a frequency 8 Hz, light intensity 200 cd/cm^2 and visual angle 2°. Each section is composed of 128×128 voxels, and for each section, 80 sample images were collected in a total of 160 s during a period of alternating between stimulation and nonstimulation block paradigms shown in Fig. 1(a). The sample time interval is 2 s.

B. Result of DSD

Similar to simulation case shown in Fig. 3, $\beta = 3$ and L =3 are chosen here. Using (7), we get the distribution of $f(p)$ and chose the voxel with intensity Z-score >0.7 (Z-score value is defined as the difference between signal and its mean value), which is obvious statistical difference approximately $\mathrm{p} < 0.001$ between the remaining voxels and the removing voxels, as the activation points as shown in Fig. 6(a). Fig. 6(b) shows the fMRI signal of the activation voxel in Fig. 6(a). The maximum correlation coefficient is 0.8402 between the activation signal in Fig. 6(b) and the experiment pattern when a sample points delay was utilized. This result confirms that the fMRI brain activation signal can be detected by DSD. The stimulus is a simple visual stimulus, and the result confirms its neural activation mainly in the occipital regions.

For comparison the feasibilities of DSD with the spatial ICA for a real fMRI data, the imaging result of spatial ICA and PCA are shown in Fig. 7. The detailed spatial ICA may be found in our prior study [12], and the PCA method is the special DSD method with $\beta = 0, L = 1$. The results of Figs. 6 and 7(a) show the result of spatial ICA and DSD are almost the same, and they are much better than the result of PCA.

V. Discussions and Conclusion

In this paper, a two steps subspace decomposition algorithm was proposed for the fMRI data processing, the first step is to reduce the noise effect on the subspace decomposition by using a delay-autocorrelation matrix instead of the general zero-DA matrix, the basis is the different smoothness of a stimulus induced fMRI signal and the intrinsic noises in a fMRI experiment. The second step is a subspace projection of the measured fMRI data. When the subspace is composed of the first principal component only, our DSD method reduces to the popular PCA

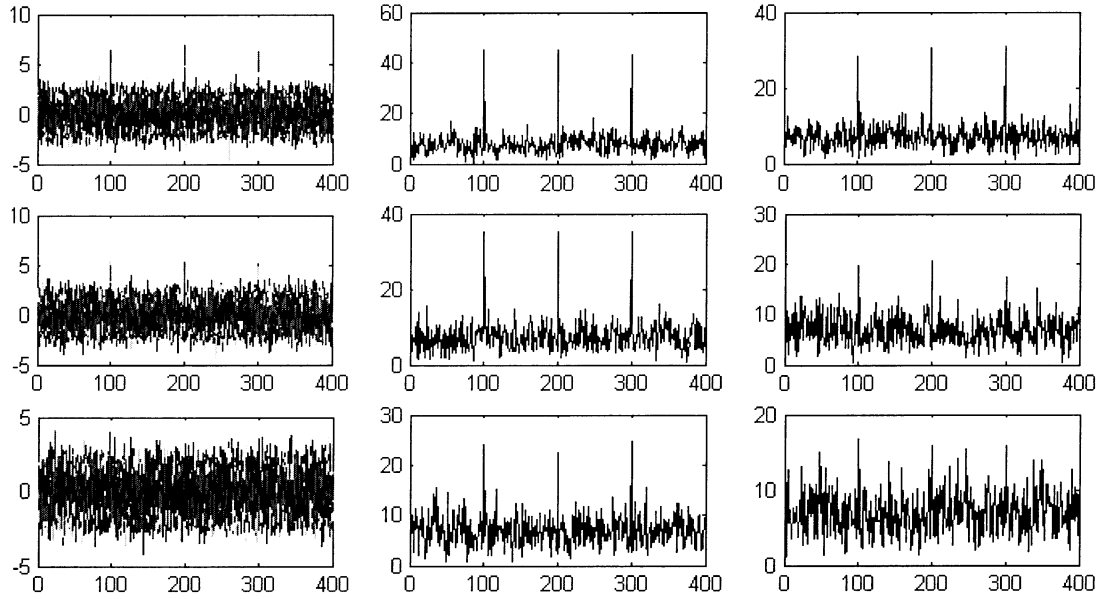

Fig. 5. Result of DSD. (ai) Simulation data with various SNR. (bi) Result of DSD with $\beta = 3$ and $L = 3$. (Ci) Result of DSD with $\beta = 0$ and $L = 1$ (PCA). The horizontal axis is the number of voxels, the vertical axis is the intensity value of the signal. Where $i = 1$ with $\mathrm{SNR} = 1.5$ dB, $i = 2$ with $\mathrm{SNR} = 1.0$ dB, and $i = 3$ with $\mathrm{SNR} = 0.7$ dB.

TABLE I
COMPARISON OF SPATIAL ICA AND DSD

Method / Activations	Spatial ICA	DSD
Synchronous	SNR>0.5 dB	SNR>0.3 dB
Asynchronous	SNR>1dB	SNR>0.6 dB

Synchronous mean the relative delay among the three activations are all zeros.

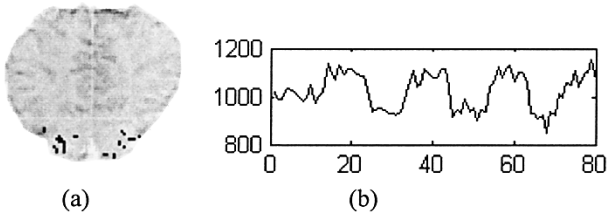

Fig. 6. fMRI by DSD. (a) fMRI result by DSD, the black points are activation voxels. (b) Activation signals. The horizontal axis is the order number of time point (one point is one second), the vertical axis is the intensity value of the signal.

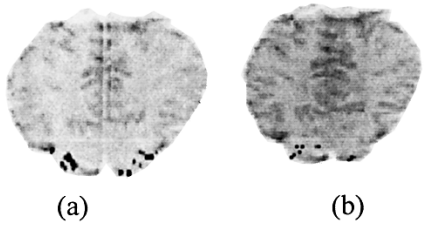

Fig. 7. fMRI by ICA and PCA. The black points are activation voxels.

method of fMRI, and it is the difference in number of principal components utilized in our DSD method and the PCA method that make our DSD method better than the PCA method. In fact, the principal components obtained by SVD are not a simple reconstruction of the real signals, thus, we cannot simply choose one of them as the most similar one of the real signal as the PCA method does, we need to utilize a combination of the main principal components as we do in our DSD method.

When comparing with the spatial ICA method, DSD and spatial ICA may produce similar results when a combination of independent components [12] is utilized as the ICA result. This fact again shows a single component no matter a principal nor an independent component cannot explain completely the fMRI signal. However, the components selection problem of spatial ICA is still a challenge topic in current practice [12], while for DSD, the problem may be the choice of the delay parameter and the truncation of the singular values for the subspace decomposition. For a simulated data, we may check the difference of DAs of the signal and the noise, and for a real fMRI data, we may check the difference of DAs of the voxel data supposed active and nonactive in a stimulus situation. Definitely, if the signal and some specific noises, such as various possible physiological pulsation noises, are similar in their statistic properties, our method may fail.

In summary, DSD is proposed to detect the regions of focal activities in the brain for fMRI data analysis. When the delay parameter $\beta = 0$, and the dimension size of the signal space $L = 1$ DSD reduce to PCA. The validity of DSD is confirmed by comparing with spatial ICA and PCA. We suppose DSD will stand as a useful tool in analyzing fMRI data in the future.

REFERENCES

[1] P. A. Bandettini, E. C. Wong, R. S. Hinks, R. S. Tikofsky, and J. S. Hyde, "Time course EPI of human brain function during task activation," *Magn. Reson. Med.*, vol. 25, pp. 390–397, 1992.

[2] P. A. Bandettini, A. Jesmanowicz, E. C. Wong, and J. S. Hyde, "Processing strategies for time-course data sets in functional MRI of the human brain," *Magn. Reson. Med.*, vol. 30, pp. 161–173, 1993.

[3] K. J. Friston, P. Jezzard, and R. Turner, "Analysis of functional MRI time-series," *Hum. Brain Mapp.*, vol. 2, pp. 69–78, 1994.

[4] F. Kruggel and D. Y. von Cramon, "Temporal properties of the hemody-namic response in functional MRI," *Hum. Brain Mapp.*, vol. 8, pp. 259–271, 1999.
[5] V. Solo, P. Purdon, R. Weisskoff, and E. Brown, "A signal estimation approach to functional MRI," *IEEE Trans. Med. Imag.*, vol. 20, no. 1, pp. 26–35, Jan. 2001.
[6] G. M. Boynton, S. A. Engel, G. H. Glover, and D. J. Heeger, "Linear systems analysis of functional magnetic resonance imaging in human V1," *J. Neurosci.*, vol. 16, no. 13, pp. 4207–4221, 1996.
[7] M. S. Cohen, "Parameteric analysis of fMRI data using linear systems methods," *NeuroImage*, vol. 6, pp. 93–103, 1997.
[8] K. J. Friston, "Statistical parametric mapping and other analyses of functional imaging data," in *Brain Mapping: The Methods*, A. W. Toga and J. C. Mazziotta, Eds. San Diego, CA: Academic, 1996, pp. 363–396.
[9] K. J. Worsley and K. J. Friston, "Analysis of fMRI time-series revisited—Again," *NeuroImage*, vol. 2, no. 1, pp. 173–181, 1995.
[10] M. J. McKeown, S. Makeig, G. G. Brown, T.-P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Analysis of fMRI data by blind separation into independent spatial components," *Hum. Brain Mapp.*, vol. 6, pp. 160–188, 1998.
[11] M. J. McKeown, T.-P. Jung, S. Makeig, G. Brown, S. S. Kindermann, T.-W. Lee, and T. J. Sejnowski, "Spatially independent activity patterns in functional MRI data during the Stroop color-naming task," *Proc. Nat. Acad. Sci.*, vol. 95, pp. 803–810, 1998.
[12] H. Chen and D. Yao, "Discussion on the choice of separation component in fMRI data analysis by spatial independent component analysis," *Magn. Reson. Imag.*, vol. 22, no. 6, pp. 827–833, 2004.
[13] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Apr. 1986.
[14] J. C. Mosher, P. S. Lewis, and R. M. Leahy, "Multiple dipole modeling and localization form spatio-temporal MEG data," *IEEE Trans. Biomed. Eng*, vol. 39, no. 6, pp. 541–557, Jun. 1992.
[15] K. Sekihara and H. Koizumi, "Detecting cortical activities from fMRI time-course data using the MUSIC algorithm with forward and backward covariance averaging," *Magn. Resource Med.*, vol. 35, no. 4, pp. 807–813, Apr. 1996.
[16] Y. W. Chiang, P. P. Borbat, and J. H. Freed, "The determination of pair distance distributions by pulsed ESR using Tikhonov regularization," *J. Magn. Reson.*, vol. 172, no. 2, pp. 279–295, Feb. 2005.
[17] R. Baumgartner, L. Ryner, W. Richter, R. Summers, M. Jarmasz, and R. Somorjai, "Comparison of two exploratory data analysis methods for fMRI: Fuzzy clustering vs. principal component analysis," *Magn. Res. Imag.*, vol. 18, pp. 89–94, 2000.
[18] H. Chen, D. Yao, Y. Zhuo, and L. Chen, "Analysis of fMRI data by blind separation into independent temporal component," *Brain Topogr.*, vol. 15, no. 4, pp. 223–232, Summer 2003.