ELSEVIER

Neurocomputing 49 (2002) 227–239

NEUROCOMPUTING

www.elsevier.com/locate/neucom

Spatio-temporal decorrelated activity patterns in functional MRI data during real and imagined motor tasks

Dawei W. Dong[a,*], J.A. Scott Kelso[a], Fred L. Steinberg[a,b]

[a] *Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL 33431, USA*

[b] *University MRI of Boca Raton, Boca Raton, FL 33431, USA*

Received 13 April 2001; accepted 22 October 2001

Abstract

We present a new method for separating multiple task-related events from other physiological and physical events revealed by functional magnetic resonance imaging (fMRI) signals. The method separates fMRI signals into different events by minimizing the spatial and temporal correlations between events. The method was used to analyze fMRI data sets from subjects performing real and imagined motor tasks. The method successfully separated task-related events (e.g, the activation in primary motor cortex during real motion and in supplementary motor area during imagined motion) from unrelated events (e.g., an activation in auditory area and slow changes probably associated with head drifts).

Keywords: Spatial–temporal decorrelation; Blind signal separation (BSS); Independent component analysis (ICA); Functional magnetic resonance imaging (FMRI)

1. Introduction

Functional magnetic resonance imaging (fMRI) data have many different signal sources, including task-related signals and signals that are not directly related to the task but correlate with head movement, breathing, etc. In the past, people have used various statistical methods to determine task-related activation in fMRI data. One of the popular methods is a Student's test, comparing the activities of an individual voxel during

* Corresponding author. Tel.: +1-561-297-2326.
E-mail address: dawei@dove.ccs.fau.edu (D.W. Dong).

0925-2312/02/$ - see front matter
PII: S0925-2312(02)00531-3

task and non-task blocks. Unfortunately, this test ignores the strong spatiotemporal statistical relationships in data [5]. To obtain a better signal-to-noise ratio and to separate different events, in recent developments, decomposition methods such as principal component analysis have been used to project data onto eigenvectors—decorrelated spatial maps [6]. This is justified if one assumes that different events are decorrelated in space and time. However, there is no guarantee that the task-related event will dominate the signal nor be one of the eigenvectors. In fact, the number of possible sets of decorrelated spatial maps is infinite. To find the correct set requires more constraints. One very interesting idea is to explore the non-Gaussian statistics of the data to separate independent signals [7,8,10]. We recently proposed a method [3] very closely related to that idea. The essential assumption is very similar: different events are unrelated or independent. However, instead of using probability distributions or higher-order spatial correlations as criteria for choosing the correct set and thus separating different events, we require that different events are spatially as well as temporally decorrelated.

The assumption is that different spatial activity patterns (events) have completely independent temporal courses. The previous method [10] assumes that the activity distribution of each event is a non-Gaussian probability density distribution. We make the assumption that the activities of different events are decorrelated and that the time-delayed activities of different events are also decorrelated [11,14], i.e., the assumption of spatio-temporal decorrelation. Here, we derive a closed-form solution that does not assume any spatial or temporal structure of an event, and works well for Gaussian and non-Gaussian distributions of different event signal sources. We show this first in a more generic fashion which uses a toy problem to illustrate the point. The method is then applied to a typical "alternate-block" paradigm used in fMRI studies in which task and rest periods alternate with each other. The method automatically separates the spatial pattern and temporal trace, activated by the main task, from other sources.

2. Method

The problem is commonly known as the mixture of sources problem. Fig. 1 shows the problem and the task of blind signal separation [1,2,9]. Independent components I_1 and I_2 are the original inputs (top two time courses on the left). S_1 and S_2 are

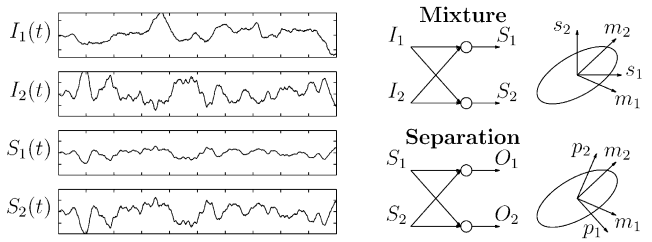

Fig. 1. Mixture of sources and blind separation.

sensor responses to the mixture of these two inputs (bottom two time courses on the left). There is a connection matrix or sensor mixture matrix which we do not know (top middle). The task is to separate the sensor responses S_1 and S_2 into the outputs O_1 and O_2 through a linear transform, so that the O_1 and O_2 are independent (bottom middle).

This is further illustrated on the right of Fig. 1 in which s_1 and s_2 are the two unit vectors (sensors) along which the response time courses S_1 and S_2 are measured, respectively. The problem is that these two sensors respond to all kinds of events. Suppose, in this example, we have two independent events—denoted by vectors m_1 and m_2—with time courses I_1 and I_2. S_1 and S_2 are combinations of I_1 and I_2 through

$$S = MI. \tag{1}$$

The mixture matrix M or the pattern of connections, with column vectors equal to m_1, m_2, and etc., is unknown to start with. The temporal traces, plotted in top two rows as I_1 and I_2, are generated, in this example, as independent Gaussian signals. They have different temporal correlation scales within themselves and there is no cross-correlation between them. Yet the responses S_1 and S_2, with some random mixture, have a common component which, in this case, resembles the inverse of I_2 because the magnitude of I_2 is 10 times larger than I_1. In other words, the mixture matrix has this specific signature because both S_1 and S_2 are dominated by I_2. The task is to find a linear transform of S_1 and S_2 that produces outputs O_1 and O_2 which are independent from each other, just as the original inputs I_1 and I_2 were independent from each other. This transform, denoted by P^{T}, gives

$$O = P^{\mathrm{T}}S, \tag{2}$$

which will equal the original input I, if $P^{\mathrm{T}} = M^{-1}$. Here P^{T} is the matrix transpose of P that is introduced for convenience since each column vector of P is then a projection vector for an independent component.

On the top right of Fig. 1, neither m_1 nor m_2 is perpendicular to the base vectors s_1 and s_2, so both I_1 and I_2 give rise to some component response in both S_1 and S_2. The joint probability density distribution of S_1 and S_2 has an elliptic shape and there are correlations between responses S_1 and S_2. Our task is to separate the output into independent components O_1 and O_2 along vectors p_1 and p_2, which are perpendicular to m_2 and m_1, respectively, with a possible permutation, where p_1 is perpendicular to m_1. It is easy to see, from the diagram on the bottom right of Fig. 1, that a solution exists.

2.1. Spatial decorrelation

In order for O_1 and O_2 to be independent they must at least be spatially decorrelated, meaning that the correlation, or covariance between O_1 and O_2 equals zero. This is the basic assumption of the second order decorrelation: in order to be independent all the correlations have to be zero. In addition, we normalize the magnitude of O_1 and O_2 so that the variance of each response equals 1. The spatial decorrelation can be

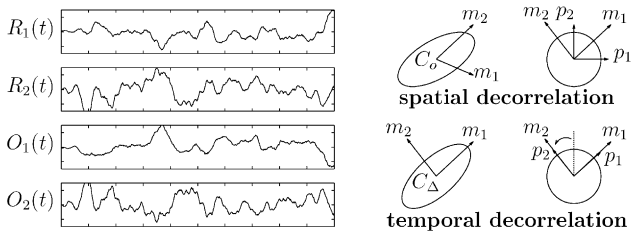

Fig. 2. Spatial and temporal decorrelation.

easily done by calculating the correlation matrix of S_1 and S_2 and then solving the eigenvectors of the correlation matrix. The decorrelation is achieved simply by

$$O = K^{\mathrm{T}} A^{-1/2} E^{\mathrm{T}} S \tag{3}$$

in which K is any orthonormal matrix, and A and E are the eigenvalue matrix and the eigenvector (column vector) matrix of the covariance matrix, i.e.,

$$\langle SS^{\mathrm{T}} \rangle$$

in which S is a column vector, $\langle\,\rangle$ denotes the average over time, and each row element of S has zero mean. This is a very standard technique that we call spatial decorrelation; it guarantees that the output is orthonormalized, i.e.,

$$\begin{aligned} &\langle O_1(t) O_2(t) \rangle = 0, \\ &\langle O_1(t) O_1(t) \rangle = 1, \\ &\langle O_2(t) O_2(t) \rangle = 1. \end{aligned} \tag{4}$$

One obvious choice of K is the identity matrix. Then the transformed signals are

$$R = A^{-1/2} E^{\mathrm{T}} S.$$

Fig. 2 (top right) shows that after the spatial decorrelation (with $K = 1$) the distribution of the output is rotationally symmetric (a sphere) when there is no correlation between the two outputs. The transform itself guarantees that. However, this decorrelation alone cannot remove all the dependence between the signals because the spatial decorrelation alone cannot uniquely determine the transform to give the independent components. As shown on Fig. 2 (top left), R_1 and R_2 traces are NOT what the original signals were (Fig. 1, top left). As one can see from the illustration in Fig. 2, one can rotate R_1 and R_2 by any angle (i.e. choose any orthonormal K) and still have the spherical distribution, which means that R_1 and R_2 are spatially decorrelated. The question then is how to determine the particular rotation that gives independent outputs. To this end, the temporal decorrelation comes into play.

2.2. Temporal decorrelation

Truly independent signals are not only decorrelated when measured at the same time, they are also decorrelated when measured with time delays. That is, the signal are not only spatially but also temporally decorrelated. The mathematical equivalence for this statement is if O_2 leads or lags O_1 by any amount of time Δ, the correlation (covariance) between these 2 variables equals zero,

$$\langle O_1(t)O_2(t+\Delta)\rangle = 0 \tag{5}$$

for a proper choice of the rotation matrix K.

Because temporal correlation exists, the joint probability density distribution C_Δ is still non-spherical for $O_1(t)$ and the time delayed $O_2(t+\Delta)$ signals, as shown on the bottom center of Fig. 2. The proper choice of K is when one finds the specific axis which points to the maximum and minimum axes. Rotating by that amount one can have p_1 and p_2 perpendicular to m_2 and m_1, respectively, or its permutation. In this way we can determine or fix the angle of rotation and obtain independent components, i.e., choose K to be the eigenvector (column vector) matrix of the delayed correlation,

$$\langle R(t)R^{\mathrm{T}}(t+\Delta)\rangle, \tag{6}$$

where, again, $\langle\,\rangle$ denotes average over time. The final transform that decorrelates the signal in both space and time is

$$P^{\mathrm{T}} = K^{\mathrm{T}}A^{-1/2}E^{\mathrm{T}}.$$

It is then straightforward to show that the original mixture matrix is

$$M = EA^{1/2}K.$$

Although E and K are both orthonormal matrices, the column vectors $m_1, m_2, \ldots,$ of the found mixture matrix $EA^{1/2}K$ (i.e., the spatial activation patterns) are not orthogonal with each other unless A is an identity matrix multiplied by a constant.

The above solution is equivalent to what was found earlier [11], although there is a major problem in real world applications, namely, how to choose the temporal scale [14]. Initially, we do not know how the signals got mixed and we do not know the time correlation scale. Ideally, if there is no noise and we have a system with two purely independent components, any chosen delay Δ will give the correct rotation. Thus, with any Δ we can obtain the temporal decorrelation rotation K and we can determine the final transform, $K^{\mathrm{T}}A^{-1/2}E^{\mathrm{T}}$. In reality, however, when we deal with data for which the temporal scale is unknown and noise is potentially involved, we may not find the same answer when we try different time delays. The temporal scale here actually plays an important role. A procedure is needed that can automatically handle this scenario. As shown in Fig. 3 there are two possibilities. Ideally, if we choose two different time delays, Δ and δ, the procedure should give two time delay distributions with the same rotation angle—maximum and minimum axes (Fig. 3, left). In reality, however, these two different time delay distributions are often not entirely coincident with each other; their maximum and minimum axes are not always the same, possibly, due to noise and non-linearity (Fig. 3, right).

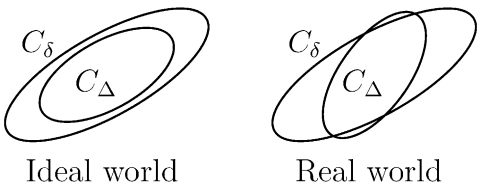

Fig. 3. Temporal scale.

In what follows, we present an analytical procedure that integrates over all possible time scales, so that we do not have to choose any particular time scale.

2.3. Temporal scale

First, let us see this problem in a more mathematical scenario. For any given time scale we have a time-delayed correlation matrix. We can represent this time-delayed correlation matrix by an eigenvector matrix multiplied by the eigenvalue matrix multiplied by the transpose of the eigenvector matrix, i.e.

$$C_\delta = K_\delta \lambda_\delta K_\delta^{\mathrm{T}}, \tag{7}$$

which is guaranteed to have everything in real numbers since we can prove straightforwardly and easily that the time-delayed correlation matrix is going to be symmetrical. On the other hand, the eigenvalues corresponding to these eigenvectors are not always positive, zero, or negative. That means we cannot just simply integrate together all those correlation matrixes corresponding to different time delays, since the negative/positive eigenvalues could cancel each other out. On the other hand, it is possible to take absolute eigenvalues individually, each to the nth power. Then we solve the eigenvector matrix of the average of those matrices, i.e.,

$$\sum_\delta K_\delta |\lambda_\delta|^n K_\delta^{\mathrm{T}}. \tag{8}$$

As a matter of fact, the absolute value of λ_δ raised to the power of n can be any positive definite operation. By doing this, the solution becomes straightforward. K—the rotation—is just the eigenvector matrix of the averaged correlation matrix. This average matrix can be calculated very straightforwardly in this equation, in particular when $n = 2$, it simply equals to $\sum_\delta C_\delta C_\delta$. We tried other monotonically increasing, non-negative functions and all worked very well.

Now the question is how well the algorithm actually performs. Let us go back to Fig. 2. The bottom left of Fig. 2 shows that after the spatial decorrelation and the rotation by the temporal decorrelation matrix the two traces O_1 and O_2 now resemble I_1 and I_2, shown in Fig. 1: one can see that O_1 has the same temporal trace as I_1 and O_2 has the same temporal trace as I_2. Note that a permutation is possible in that O_1 and O_2 can correspond to I_1 and I_2, but can also correspond to I_2 and I_1. There is also another degree of freedom besides permutation, the scale. O_1 and O_2 can be on different scales from I_1 and I_2. But the important point is the successful separation of

the output into two channels which are, after permutation and scale, the same as the two input channels.

3. Experiment

Now let us examine what happens when we apply this algorithm to real data from a realistic problem. In this case we use brain image data collected using fMRI.

3.1. fMRI experimental setup

Anatomical and functional magnetic resonance images were acquired using a 1.5 T G.E. Signa NV/i MR scanner (General Electric Medical Systems, Milwaukee, WI). At the beginning of each experimental session, anatomical images of 256×256 spatial resolution were collected for 20 axial slices, using a 2D gradient echo SPGR pulse sequence with the parameters TE = in phase (4.6 ms), TR = 325 ms, flip angle = 90°, bandwidth 15.63 KHz. The field of view was 240 mm. The axial slices were 4.0 mm thick with 2.0 mm gap. These images were later used to overlay fMRI activations on the anatomical structure. During the experiment, functional images of 64×64 spatial resolution were acquired at the same 20 slice locations, using a gradient echo EPI sequence with the parameters TE = 60 ms, TR = 3000 ms, flip angle = 90°, bandwidth = 62.5 KHz. Thus each functional scan has $64 \times 64 \times 20$ voxels of dimension 3.75 mm $\times$ 3.75 mm $\times$ 6.0 mm.

A typical "alternate-block" task was used. In the first block the subject was asked to perform a certain task, in this particular case—to tap his/her finger for 30 s. In the next block the subject rests without moving. Subsequent blocks alternate finger tapping and resting, for a total of 10 blocks of 30 s each. Since the whole brain image was scanned every 3 s, there are a total of 100 samples in time. Fig. 4 (top left) illustrates the time series of this task lasting 300 s, with 5 *on* and 5 *off* periods (blocks). Three different ways of finger tapping are used: (1) simple flexion performed with index finger of one hand flexing and extending, (2) sensorimotor sequences with thumb touching other

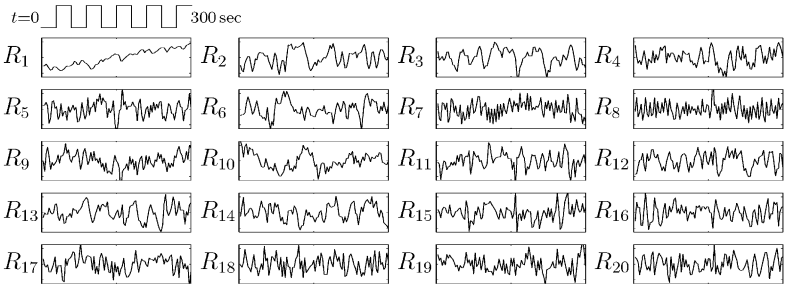

Fig. 4. Spatial decorrelation. Shown in the figure are the top 20 temporal traces which are decorrelated spatially. The zero of vertical axis is at the middle of each plot. The scale of the vertical axis is not marked, since only the relative amplitude and the relative sign are meaningful.

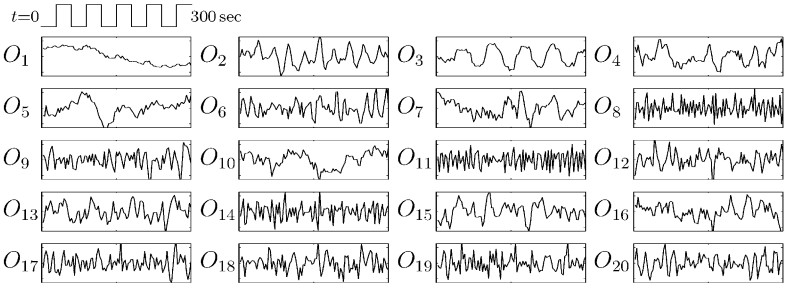

Fig. 5. Temporal decorrelation. Shown in the figure are the top 20 temporal traces which are decorrelated spatially and temporally. The zero of vertical axis is at the middle of each plot with upward for positive and downward for negative (see the text for the sign convention). For the top three temporal traces O_1, O_2, and O_3, the corresponding activity patterns are shown in Fig. 6.

fingers of the same hand in given sequences, (3) imagining sensorimotor sequences as in (2) without actually moving any finger.

3.2. Decorrelated time series

In this experiment, we have a sensor array of $64 \times 64 \times 20$ voxels. We can calculate correlations between those voxels during task performance. Then we can solve the correlation matrix and look at the different eigenvector responses, i.e., temporal responses of spatially decorrelated channels. Fig. 4 shows the 20 eigenvectors (time series) with the largest eigenvalues (i.e., the greatest signal powers). The eigenvectors and eigenvalues are solved using the singular value decomposition method [12]. The eigenvector in the top left graph (R_1) has the largest eigenvalue and the eigenvalues decrease from left to right and top to bottom. In fact, the top 20 traces account for 70% of the signal variance. The lowest traces are mostly experimental noise. The noise variance is estimated to be equal to the lowest non-zero eigenvalue. After the singular value decomposition, we keep only the top 20 eigenvectors, which are the ones with eigenvalues at least two times larger than the estimated noise variance, i.e., with a signal-to-noise ratio larger than one.

The task time series is superimposed on the top left eigenvector. As one can see, none of the 20 temporal traces are related to the task in an obvious way. The next step, then, is to apply the described temporal decorrelation to eliminate the arbitrariness in the rotation and fix the real transformation which will result in independent components. The result is shown in Fig. 5. As one can see the third temporal trace (the top trace in the third column, O_3) closely resembles the task block, with *on* and *off* periods that correspond well with the task timing.

3.3. Activity patterns

In Fig. 6 the brain activity patterns of the top three components are superimposed on the gray-scale high resolution anatomical images. In each plot, only ten cortical

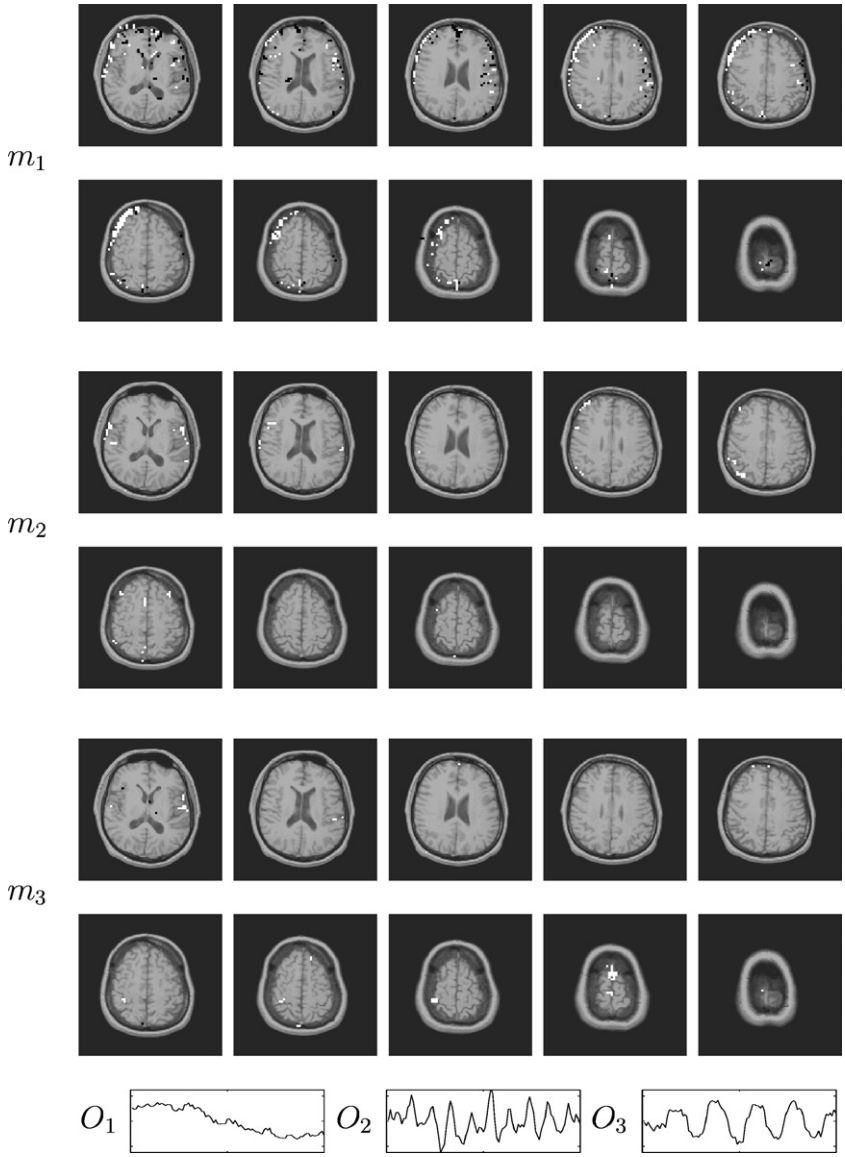

Fig. 6. fMRI activity patterns during simple flexion performed with the left hand of a subject. Shown on the top are the leading three activity patterns m_1, m_2, and m_3, for which the corresponding temporal traces O_1, O_2, and O_3 are shown at the bottom (the same as the leading ones in Fig. 5). They account for 17.7%, 4.7%, and 4.4% of the signal variance.

slices are shown. When plotting each slice, the radiology convention is followed, i.e., anterior—top, posterior—bottom, left ear—right, and right ear—left. For each spatial–temporal component, the signal variance for each voxel is calculated. A voxel is plotted

in white for positive and black for negative,[1] only when the calculated variance is at least nine times larger than the estimated noise variance.

We observe a task-related pattern (Fig. 6, m_3) for which the main activation is in the motor cortex with some activation in the supplementary motor area (SMA) and in the primary auditory area. The corresponding temporal trace is shown in Figs. 5 and 6, O_3.

There is another task-related pattern (Fig. 6, m_2) but the activation in this case is at the onset and offset of each task block (Figs. 5 and 6, O_2). In this case, we see bilateral auditory area activation and activation in other brain areas. The bilateral auditory activation may stem from the oral instructions given at the beginning of each block: "start to tap your fingers" or "stop tapping your fingers". Also, the subject probably anticipates the task for the upcoming 30 s, reflected by activity in other brain areas.

There is also a slow varying component (Figs. 5 and 6, O_1), probably corresponding to the change of position in the head, likely because we did not restrain the subject with any fixational device, like a bite-bar, so the subject's head during a trial of 300 s may have moved from one position to another. Its particular spatial pattern (Fig. 6, m_1) is consistent with movements along the ear-to-ear axis. There are other slow varying components, possibly corresponding to movements along different axes. Of course, this is our interpretation of these activations but it has been similarly interpreted by other groups [10]. The rest of the components are much harder to interpret and most likely contain more noise.

In summary, by applying this method we are able to separate the task related event in a variety of different finger movement conditions. For simple flexion, the method showed activation in contralateral motor/somatosensory cortexes and bilateral SMA. For sensorimotor sequences, it showed the activation in contralateral motor/somatosensory cortexes and bilateral SMA, premotor, prefrontal, and parietal cortexes. For imagining sensorimotor sequences, the method revealed activation in bilateral SMA, premotor, and prefrontal cortexes. Our results (Fig. 7 shows three of the trials) are consistent with studies using other methods (see for example [13]). Although, we demonstrated how the method can separate the signals due to head movements, we are not necessarily advocating that one relies on it. In fact, the head should be stabilized mechanically whenever possible and the proven method of registering serial 3D fMRI should be used [4]. But after all these are done, the method may be still needed to separate different sources of signal and noise.

4. Conclusion

We found a method of source separation by spatio-temporal decorrelation. This algorithm runs very fast (less than a minute on a 550 MHz Pentium^{TM} PC). This is

[1] The sign convention for a spatial–temporal component, i.e., a pair of temporal trace O and activity pattern m, is the following: at a given time t and a given voxel location x, when the value $O(t)$ of the temporal trace and the value $m(x)$ of the activity pattern are both positive or both negative, it means an increase in the blood oxygen level dependent ($BOLD$) contrast; when one is positive and the other is negative, it means a decrease in $BOLD$ contrast.

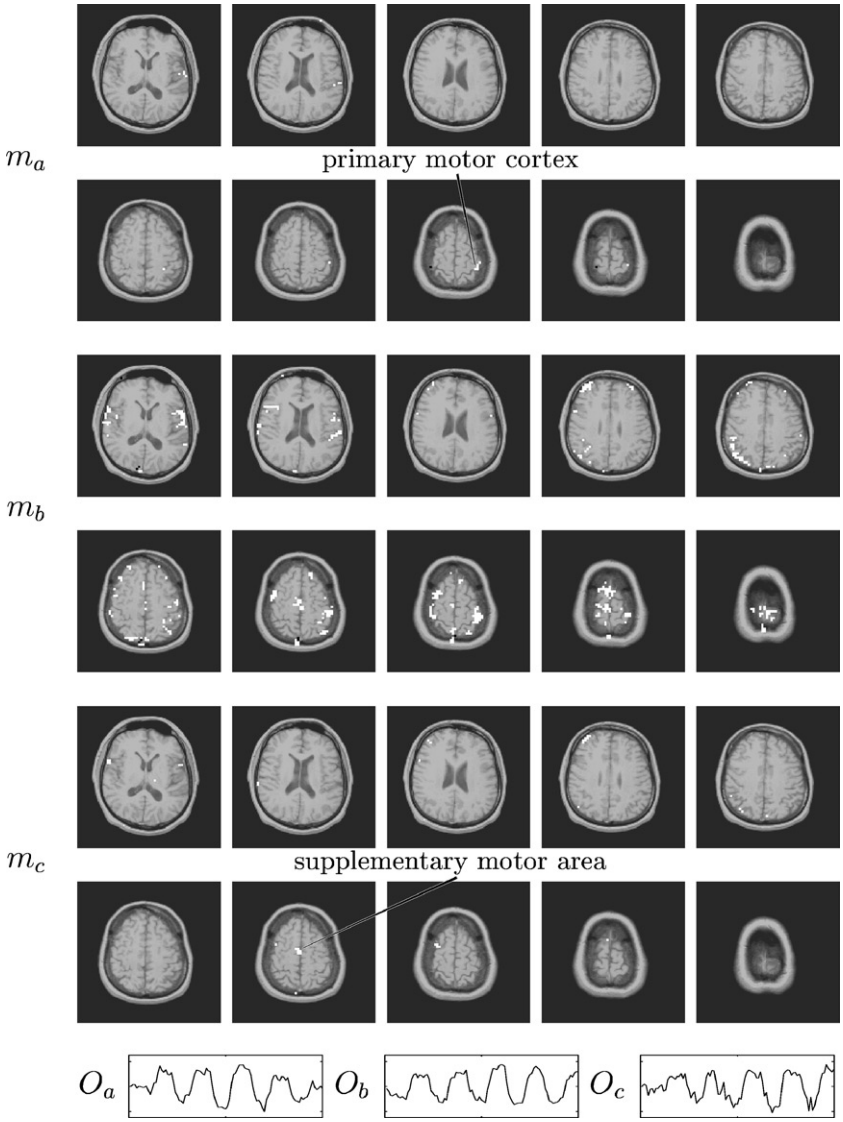

Fig. 7. fMRI activity patterns for three different ways of finger tapping performed with the right hand of a subject. They are (a) simple flexion, (b) sensorimotor sequences, and (c) imagining sensorimotor sequences. Shown on the top are the three task-related activity patterns m_a, m_b, and m_c, for which the corresponding temporal traces O_a, O_b, and O_c are shown at the bottom. They account for 3.0%, 12.2%, and 2.2% of the signal variances of the corresponding experiments, respectively.

much faster than the independent component analysis [10]. We are able to separate task-related from task-unrelated events which gives us better signal-to-noise ratio in studying brain function. This allows us to minimize the effects of head movement and

other physical or physiological events unrelated to the tasks the subject is performing. Also, and this has not, to our knowledge, been reported before, the algorithm is not limited to identifying a single task-related event but can separate multiple task-related brain activities and their time courses. As demonstrated here, it can separate task-related events which are spatially and temporally uncorrelated. Our method is an analytical method. It works in a scenario with no prior assumptions about the temporal scale or about the probability distribution (i.e., Gaussian or non-Gaussian) of different events.

Supplementary information

It is available on the authors' web site (http://dove.ccs.fau.edu/01-MRI-add.html).

Acknowledgements

We thank Dale Wilke for his excellent technical support. DWD is especially indebted to Dale Wilke for helping to run the MRI experiments late at night and in a few cases on overtime. We wish to acknowledge Xin Guan, Makoto Fukushima, Justine Mayville, and Armin Fuchs for their assistance in data collection and Anna Kashina for her assistance in writing an earlier draft of the paper. We thank Betty Tuller and two anonymous reviewers for their helpful comments on the manuscript. This work is partially supported by NIMH grant MH42900.

References

[1] S.I. Amari, A. Cichocki, Adaptive blind signal processing—neural network approaches, Proc. IEEE 86 (10) (1998) 2026–2048.
[2] U.M. Bae, T.W. Lee, S.Y. Lee, Blind signal separation in teleconferencing using ICA mixture model, Electron. Lett. 36 (7) (2000) 680–682.
[3] D.W. Dong, J.A.S. Kelso, W.D. Wilke, F. Steinberg, Spatiotemporal decorrelated activity patterns in functional MRI data during real and imagery motor tasks, Abstr. Soc. Neurosci. 25 (1999) 786.
[4] P.A. Freeborough, R.P. Woods, N.C. Fox, Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders, J. Comput. Assist. Tomo. 20 (6) (1996) 1012–1022.
[5] K.J. Friston, Statistical parametric mapping and other analyses of functional imaging data, in: A.W. Toga, J.C. Mazziotta (Eds.), Brain Mapping, The Methods, Academic Press, San Diego, CA, 1996, pp. 363–396.
[6] K.J. Friston, C.D. Frith, P.F. Liddle, R.S.J. Frackowiak, Functional connectivity—the principal-component analysis of large (PET) data sets, J. Cereb. Blood Flow Metab. 13 (1) (1993) 5–14.
[7] J.J. Hopfield, Olfactory computation and object perception, Proc. Nat. Acad. Sci. USA 88 (15) (1991) 6462–6466.
[8] C. Jutten, J. Herault, A. Guerin, Artificial Intelligence and Cognitive Sciences, Manchester University Press, Manchester, 1988, pp. 231–248.
[9] S. Makeig, T.P. Jung, A.J. Bell, D. Ghahremani, T.J. Sejnowski, Blind separation of auditory event-related brain responses into independent components, Proc. Nat. Acad. Sci. USA 94 (20) (1997) 10979–10984.

[10] M.J. McKeown, T.P. Jung, S. Makeig, G. Brown, S.S. Kindermann, T.W. Lee, T.J. Sejnowski, Spatially independent activity patterns in functional MRI data during the Stroop color-naming task, Proc. Nat. Acad. Sci. USA 95 (3) (1998) 803–810.
[11] L. Molgedey, H.G. Schuster, Separation of a mixture of independent signals using time-delayed correlations, Phys. Rev. Lett. 72 (23) (1994) 3634–3637.
[12] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes: The Art of Scientific Computing, 2nd Edition, Cambridge University Press, Cambridge, 1992.
[13] P.E. Roland, Brain Activation, Wiley-Liss, New York, 1993.
[14] H.C. Wu, J.C. Principe, A unifying criterion for blind source separation and decorrelation: simultaneous diagonalization of correlation matrices, in: J. Principe, L. Giles, N. Morgan, E. Wilson (Eds.), Neural Networks for Signal Processing, Vol. VII, IEEE, Piscataway, NJ, 1997, pp. 496–505.

Dawei W. Dong is an assistant professor at the Center for Complex Systems and Brain Sciences at Florida Atlantic University. His research goal is to uncover fundamental principles of how the nervous system codes and uses sensory information.

J.A. Scott Kelso holds the Glenwood and Martha Creech Chair of Science and is the director of The Center for Complex Systems and Brain Sciences at Florida Atlantic University. His goal is to understand the nature of coordination within and among areas of the human brain and their relation to behavior. His research embraces experimentation, analysis and theoretical modeling at both behavioral and neural levels using large scale EEG, MEG and functional Magnetic Resonance Imaging methods.

Fred L. Steinberg, M.D. is a radiologist and the Medical Director of University MRI of Boca Raton. He is also a research professor at the Center for Complex Systems and Brain Sciences at Florida Atlantic University.