

A SPATIALLY ROBUST ICA ALGORITHM FOR MULTIPLE fMRI DATA SETS

Ana S. Lukic,¹ Miles N. Wernick,¹ Lars Kai Hansen,² Jon Anderson,³ Stephen C. Strother³

¹Illinois Institute of Technology, Chicago, ²Technical University of Denmark, Lyngby, Denmark

³University of Minnesota and VA Medical Center, Minneapolis

ABSTRACT

In this paper we derive an independent-component analysis (ICA) method for analyzing two or more data sets simultaneously. Our model extracts independent components common to all data sets and independent data-set-specific components. We use time-delayed autocorrelations to obtain independent signal components and base our algorithm on prediction analysis. We applied this method to functional brain mapping using functional magnetic resonance imaging (fMRI). The results of our 3-subject analysis demonstrate the robustness of the algorithm to the spatial misalignment intrinsic in multiple-subject fMRI data sets.

1. INTRODUCTION

There are several approaches to the problem of blind separation of linear signal mixtures. Some of them are based on non-Gaussianity [1, 3] while others exploit temporal correlations [2, 6]. The technique of Molgedey and Schuster [6] is especially attractive since it offers a non-iterative solution. The limitations of the original algorithm, such as restriction to square non-singular mixing matrices and inherent erroneous complex-valued results, were eliminated in [5]. Different ICA techniques are applied to fMRI and compared in [7].

In this paper, we extend the Molgedey-Schuster ICA method to analyze more than one data set simultaneously. There are many potential applications for simultaneous independent-component analysis (ICA) of multiple data sets. For example, one might simultaneously analyze the audio and video portions of a video sequence.

Our interest is in developing methods to analyze functional magnetic resonance images (fMRI) of the brain, to map spatial and temporal patterns of brain activation. In this application, the data can be divided into multiple data sets by considering different runs within a subject and/or different subjects separately, or by considering image data separately from other measurements made during the fMRI study (e.g., reaction time of the subject). In this paper we illustrate the first approach.

2. EXTENDED BLIND SIGNAL SEPARATION PROBLEM

In the classical blind signal separation problem it is assumed that there are N independent source signals in matrix \mathbf{S} observed through a mixing matrix \mathbf{A} , i.e.,

$$\mathbf{X} = \mathbf{A}\mathbf{S}, \quad (1)$$

to obtain a matrix of observations \mathbf{X} . The goal in this classical ICA problem is to recover the sources in \mathbf{S} and the mixing matrix \mathbf{A} without any prior knowledge except the assumption of independence of the sources in \mathbf{S} .

To analyze two data sets \mathbf{X} and \mathbf{Y} simultaneously, we could lump them together into one matrix, and use the model in (1); however, this would force the description of the two matrices to be based on the same set of sources. To allow the two data sets to have some common components, and other data-set-specific components, we propose the following model:

$$\begin{aligned} \mathbf{X} &= \mathbf{A}_x \mathbf{S}_x + \mathbf{B}_x \mathbf{T} \\ \mathbf{Y} &= \mathbf{A}_y \mathbf{S}_y + \mathbf{B}_y \mathbf{T} \end{aligned} \quad (2)$$

where \mathbf{X} and \mathbf{Y} are two sets of observations with each row corresponding to a voxel time series at a given spatial location in the brain; \mathbf{S}_x and \mathbf{S}_y are independent sources specific to \mathbf{X} and \mathbf{Y} , respectively; \mathbf{T} is a matrix whose rows are K independent time-sources common to both \mathbf{X} and \mathbf{Y} ; \mathbf{A}_x , \mathbf{A}_y , \mathbf{B}_x , and \mathbf{B}_y are mixing matrices whose columns are the images that correspond to the extracted independent components. The numbers of voxels in \mathbf{X} and \mathbf{Y} are denoted by L and P , respectively; the numbers of sources in \mathbf{S}_x and \mathbf{S}_y are N and M , respectively. We will describe the algorithm in terms of only two data sets, but it is easily extended to more than two.

3. MIXING MATRICES AS EIGENVECTOR MATRICES

In this section we show that the mixing matrices in this ICA problem can be obtained as eigenvectors of linear minimum mean-square error (LMMSE) prediction matrices.

Assuming the independent signals have non-vanishing time autocorrelation functions we suppose there is a LMMSE prediction matrix \mathbf{W} such that:

$$\begin{bmatrix} \mathbf{X}_\tau \\ \mathbf{Y}_\tau \end{bmatrix} = \underset{((L+P) \times (L+P))}{\mathbf{W}} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \quad (3)$$

where \mathbf{X}_τ and \mathbf{Y}_τ are delayed versions of \mathbf{X} and \mathbf{Y} respectively. By combining (2) and (3) we can write:

$$\begin{bmatrix} \mathbf{A}_x \mathbf{S}_{x\tau} + \mathbf{B}_x \mathbf{T}_\tau \\ \mathbf{A}_y \mathbf{S}_{y\tau} + \mathbf{B}_y \mathbf{T}_\tau \end{bmatrix} = \begin{bmatrix} \underset{(L \times L)}{\mathbf{W}_{xx}} & \underset{(L \times P)}{\mathbf{W}_{xy}} \\ \underset{(P \times L)}{\mathbf{W}_{yx}} & \underset{(P \times P)}{\mathbf{W}_{yy}} \end{bmatrix} \begin{bmatrix} \mathbf{A}_x \mathbf{S}_x + \mathbf{B}_x \mathbf{T} \\ \mathbf{A}_y \mathbf{S}_y + \mathbf{B}_y \mathbf{T} \end{bmatrix} \quad (4)$$

where we have divided the matrix \mathbf{W} into four blocks \mathbf{W}_{xx} , \mathbf{W}_{xy} , \mathbf{W}_{yx} and \mathbf{W}_{yy} ; $\mathbf{S}_{x\tau}$, $\mathbf{S}_{y\tau}$ and \mathbf{T}_τ are delayed versions of \mathbf{S}_x , \mathbf{S}_y and \mathbf{T} , respectively.

Right multiplying (4) by \mathbf{S}_x^T yields:

$$\begin{bmatrix} \mathbf{A}_x \mathbf{S}_{x\tau} \mathbf{S}_x^T + \mathbf{B}_x \mathbf{T}_\tau \mathbf{S}_x^T \\ \mathbf{A}_y \mathbf{S}_{y\tau} \mathbf{S}_x^T + \mathbf{B}_y \mathbf{T}_\tau \mathbf{S}_x^T \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{xx} & \mathbf{W}_{xy} \\ \mathbf{W}_{yx} & \mathbf{W}_{yy} \end{bmatrix} \begin{bmatrix} \mathbf{A}_x \mathbf{S}_x \mathbf{S}_x^T + \mathbf{B}_x \mathbf{T} \mathbf{S}_x^T \\ \mathbf{A}_y \mathbf{S}_y \mathbf{S}_x^T + \mathbf{B}_y \mathbf{T} \mathbf{S}_x^T \end{bmatrix} \quad (5)$$

Because the sources in \mathbf{S}_x , \mathbf{S}_y and \mathbf{T} are independent, $\mathbf{S}_{y\tau} \mathbf{S}_x^T = 0$, $\mathbf{S}_y \mathbf{S}_x^T = 0$, $\mathbf{T}_\tau \mathbf{S}_x^T = 0$ and $\mathbf{T} \mathbf{S}_x^T = 0$. Thus, (5) becomes:

$$\begin{bmatrix} \mathbf{A}_x \mathbf{C}_x(\tau) \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{xx} & \mathbf{W}_{xy} \\ \mathbf{W}_{yx} & \mathbf{W}_{yy} \end{bmatrix} \begin{bmatrix} \mathbf{A}_x \mathbf{C}_x(0) \\ 0 \end{bmatrix}, \quad (6)$$

where $\mathbf{C}_x(\tau) = \mathbf{S}_{x\tau} \mathbf{S}_x^T$ and $\mathbf{C}_x(0) = \mathbf{S}_x \mathbf{S}_x^T$. The first row of (6) yields

$$\mathbf{W}_{xx} \mathbf{A}_x = \mathbf{A}_x [\mathbf{C}_x(\tau) \mathbf{C}_x(0)^{-1}] \quad (7)$$

which is an eigenvector equation for \mathbf{W}_{xx} since $\mathbf{C}_x(\tau) \mathbf{C}_x(0)^{-1}$ is diagonal. The second row of (6) yields

$$0 = \mathbf{W}_{yx} \mathbf{A}_x \quad (8)$$

since $\mathbf{C}_x(0)$ is a full rank positive diagonal matrix.

Similarly, by right-multiplying (4) by \mathbf{S}_y^T we obtain:

$$0 = \mathbf{W}_{xy} \mathbf{A}_y \mathbf{C}_y(0) \quad (9)$$

$$\mathbf{W}_{yy} \mathbf{A}_y = \mathbf{A}_y [\mathbf{C}_y(\tau) \mathbf{C}_y(0)^{-1}] \quad (10)$$

Finally, after right multiplying (4) by \mathbf{T}^T we obtain:

$$\begin{bmatrix} \mathbf{B}_x \\ \mathbf{B}_y \end{bmatrix} \mathbf{C}_T(\tau) = \begin{bmatrix} \mathbf{W}_{xx} & \mathbf{W}_{xy} \\ \mathbf{W}_{yx} & \mathbf{W}_{yy} \end{bmatrix} \begin{bmatrix} \mathbf{B}_x \\ \mathbf{B}_y \end{bmatrix} \mathbf{C}_T(0) \quad (11)$$

or, rearranging,

$$\mathbf{W} \begin{bmatrix} \mathbf{B}_x \\ \mathbf{B}_y \end{bmatrix} = \begin{bmatrix} \mathbf{B}_x \\ \mathbf{B}_y \end{bmatrix} [\mathbf{C}_T(\tau) \mathbf{C}_T(0)^{-1}] \quad (12)$$

Equations (7), (10), and (12) are eigenvector equations that provide a recipe for identifying the various mixing matrices as eigenvectors of prediction matrices [5].

4. THE ALGORITHM

We divide the algorithm in two parts. First, we estimate the prediction matrix \mathbf{W} and then we use it to estimate mixing matrices and independent sources. The LMMSE prediction matrix is [8]:

$$\hat{\mathbf{W}}(\tau) = \hat{\mathbf{R}}_{\mathbf{X}\mathbf{Y}}(\tau) \hat{\mathbf{R}}_{\mathbf{X}\mathbf{Y}}(0)^{-1}, \quad (13)$$

where $\hat{\mathbf{R}}_{\mathbf{X}\mathbf{Y}}(0)$ is an estimate of the autocorrelation function at lag 0,

$$\hat{\mathbf{R}}_{\mathbf{X}\mathbf{Y}}(0) = \left(\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \begin{bmatrix} \mathbf{X}^T & \mathbf{Y}^T \end{bmatrix} \right), \quad (14)$$

and $\hat{\mathbf{R}}_{\mathbf{X}\mathbf{Y}}(\tau)$ is an estimate of the autocorrelation function at lag τ ,

$$\hat{\mathbf{R}}_{\mathbf{X}\mathbf{Y}}(\tau) = \frac{1}{2} \left(\begin{bmatrix} \mathbf{X}_\tau \\ \mathbf{Y}_\tau \end{bmatrix} \begin{bmatrix} \mathbf{X}^T & \mathbf{Y}^T \end{bmatrix} + \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \begin{bmatrix} \mathbf{X}_\tau^T & \mathbf{Y}_\tau^T \end{bmatrix} \right) \quad (15)$$

where \mathbf{X}_τ and \mathbf{Y}_τ are cyclic permutations by τ steps, *i.e.* $\mathbf{X}_\tau = \{\mathbf{x}((n+\tau)_T)\}$, $\mathbf{x}(k)$ denotes the k^{th} column of \mathbf{X} and $(\cdot)_T$ denotes the argument modulo T . Equation (15) guarantees that $\hat{\mathbf{R}}_{\mathbf{X}\mathbf{Y}}(\tau)$ is symmetric with real eigenvalues [5]. Consider the singular value decompositions (SVDs):

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \mathbf{U} \mathbf{D} \mathbf{V}^T, \quad \begin{bmatrix} \mathbf{X}_\tau \\ \mathbf{Y}_\tau \end{bmatrix} = \mathbf{U}_\tau \mathbf{D}_\tau \mathbf{V}_\tau^T \quad (16)$$

where eigenvector matrices contain only components whose corresponding eigenvectors are nonzero. Since $\mathbf{X}\mathbf{X}^T = \mathbf{X}_\tau \mathbf{X}_\tau^T$ it follows that $\mathbf{U}_\tau = \mathbf{U}$ and $\mathbf{D}_\tau = \mathbf{D}$. Then:

$$\begin{aligned} \hat{\mathbf{W}}(\tau) &= \frac{1}{2} \left(\mathbf{U} \mathbf{D} \mathbf{V}_\tau^T (\mathbf{U} \mathbf{D} \mathbf{V}^T)^T + \mathbf{U} \mathbf{D} \mathbf{V}^T (\mathbf{U} \mathbf{D} \mathbf{V}_\tau^T)^T \right) (\mathbf{U} \mathbf{D} \mathbf{V}^T (\mathbf{U} \mathbf{D} \mathbf{V}^T)^T)^{-1} \\ &= \frac{1}{2} \mathbf{U} \mathbf{D} (\mathbf{V}_\tau^T \mathbf{V} + \mathbf{V}^T \mathbf{V}_\tau) \mathbf{D}^{-1} \mathbf{U}^T \end{aligned} \quad (17)$$

We will estimate $\hat{\mathbf{W}}$ as the average of $\hat{\mathbf{W}}(\tau)$ over a set of values for τ :

$$\hat{\mathbf{W}} = \frac{1}{t} \sum_{j=1}^t \hat{\mathbf{W}}(\tau_j) = \frac{1}{2t} \mathbf{U} \mathbf{D} \left(\sum_{j=1}^t (\mathbf{V}_{\tau_j}^T \mathbf{V} + \mathbf{V}^T \mathbf{V}_{\tau_j}) \right) \mathbf{D}^{-1} \mathbf{U}^T \quad (18)$$

where t is the number of different values of τ we want to average over and $\tau_j, j=1 \dots t$ are those values. We are now ready to outline the algorithm:

PART I: *Calculating the prediction matrix $\hat{\mathbf{W}}$*

Step 1: Perform the SVD in (16), keep only the components whose eigenvalue are nonzero as a denoising step.

Step 2: Calculate $\mathbf{K} = \sum_{j=1}^t (\mathbf{V}_{\tau_j}^T \mathbf{V} + \mathbf{V}^T \mathbf{V}_{\tau_j})$,

Step 3: Calculate $\hat{\mathbf{W}} = \frac{1}{2t} \mathbf{U} \mathbf{D} \mathbf{K} \mathbf{D}^{-1} \mathbf{U}^T$

PART II: Estimating the mixing matrices

Step 1: Estimate the columns of $\hat{\mathbf{A}}_x$ as the eigenvectors of \mathbf{W}_{xx} but select only those that satisfy (8) and have a corresponding nonzero eigenvalue. Denote the number of such eigenvectors (columns of $\hat{\mathbf{A}}_x$) by N .

Step 2: Estimate the columns of $\hat{\mathbf{A}}_y$ as the eigenvectors of \mathbf{W}_{yy} but select only those that satisfy (9) and have a corresponding nonzero eigenvalue. Denote the number of such eigenvectors (columns of $\hat{\mathbf{A}}_y$) by M .

Step 3: Estimate $\hat{\mathbf{B}} = [\hat{\mathbf{B}}_x^T \ \hat{\mathbf{B}}_y^T]^T$ as the matrix of K eigenvectors of \mathbf{W} corresponding to the K largest eigenvalues of \mathbf{W} where: $K = \text{rank}([\mathbf{X}^T \ \mathbf{Y}^T]) - N - M$.

PART III: Estimating the independent sources

Step 1: From $\mathbf{X} = [\mathbf{A}_x \ \mathbf{B}_x] \begin{bmatrix} \mathbf{S}_x \\ \mathbf{T} \end{bmatrix}$ calculate:

$$\begin{bmatrix} \hat{\mathbf{S}}_x \\ \hat{\mathbf{T}}_x \end{bmatrix} = \left(\begin{bmatrix} \hat{\mathbf{A}}_x^T \\ \hat{\mathbf{B}}_x^T \end{bmatrix} \begin{bmatrix} \hat{\mathbf{A}}_x & \hat{\mathbf{B}}_x \end{bmatrix} \right)^{-1} \begin{bmatrix} \hat{\mathbf{A}}_x^T \\ \hat{\mathbf{B}}_x^T \end{bmatrix} \mathbf{X} \quad (19)$$

Step 2: Calculate:

$$\begin{bmatrix} \hat{\mathbf{S}}_y \\ \hat{\mathbf{T}}_y \end{bmatrix} = \left(\begin{bmatrix} \hat{\mathbf{A}}_y^T \\ \hat{\mathbf{B}}_y^T \end{bmatrix} \begin{bmatrix} \hat{\mathbf{A}}_y & \hat{\mathbf{B}}_y \end{bmatrix} \right)^{-1} \begin{bmatrix} \hat{\mathbf{A}}_y^T \\ \hat{\mathbf{B}}_y^T \end{bmatrix} \mathbf{Y} \quad (20)$$

Step 3: Calculate $\hat{\mathbf{T}} = \frac{\hat{\mathbf{T}}_x + \hat{\mathbf{T}}_y}{2}$, since \mathbf{T}_x and \mathbf{T}_y should be identical.

5. APPLICATION

We applied the algorithm to a 3-subject set of fMRI images. Three right-handed volunteers performed two runs of a static force task [4] alternating six rest and five force periods/run (44 s/period; 200, 400, 600, 800, 1000 g force levels between thumb and forefinger, pseudo-randomized across force periods and maintained with visual feedback). Data were collected with a Siemens 1.5T clinical scanner (fMRI: EPI BOLD, TR/TE=3986/60 msec, FOV=22x22x15 cm, slices=30, voxel=3.44x3.44x5 mm; MRI: T1-weighted 3D FLASH) and preprocessed as follows: (1) spatial smoothing/slice with a 2D Gaussian (FWHM = 2.0 pixels); (2) using AIR 3.08 for within-subject registration (six-parameter) followed by between-subject registration (12-parameter affine) to a reference MRI in Talairach space; (3) temporal detrending with a 0.5 cycle cosine basis function. Before analyzing the data sets 2-3 transition scans that reflect the hemodynamic response were dropped from the beginning and end of each period.

Initially each subject was processed individually. First, principal component analysis was performed and the data were denoised by keeping only the first 50 principal components. The described ICA analysis was then performed to locate the task-sensitive brain regions and common time sequences in both runs from each

subject. To do that we constructed two data matrices \mathbf{X}_i and \mathbf{Y}_i corresponding to the first and second runs of subject i . Each row of \mathbf{X}_i and \mathbf{Y}_i corresponds to a voxel and each column is one scan in time. We then further denoised each single-subject data set by keeping only the components common to both runs:

$$\begin{bmatrix} \hat{\mathbf{X}}_i \\ \hat{\mathbf{Y}}_i \end{bmatrix} = \begin{bmatrix} \mathbf{B}_{x_i} \tilde{\mathbf{T}}_i \\ \mathbf{B}_{y_i} \tilde{\mathbf{T}}_i \end{bmatrix} \quad (21)$$

where $\tilde{\mathbf{T}}_i$ is obtained from \mathbf{T}_i by keeping the 10 time sequences that best correlate with the known stimulus. We then analyzed the denoised data from all three subjects using an extension of the algorithm applied simultaneously to the three sets of subject's observations:

$$\begin{aligned} \mathbf{D}_1 &= \mathbf{A}_1 \mathbf{S}_1 + \mathbf{B}_1 \mathbf{T} \\ \mathbf{D}_2 &= \mathbf{A}_2 \mathbf{S}_2 + \mathbf{B}_2 \mathbf{T} \\ \mathbf{D}_3 &= \mathbf{A}_3 \mathbf{S}_3 + \mathbf{B}_3 \mathbf{T} \end{aligned} \quad (22)$$

where $\mathbf{D}_i = [\hat{\mathbf{X}}_i \ \hat{\mathbf{Y}}_i]$. The algorithm extracted seven time sequences common to all three subjects. Figure 1 shows the one that correlates the best with the known stimulus (dashed line).

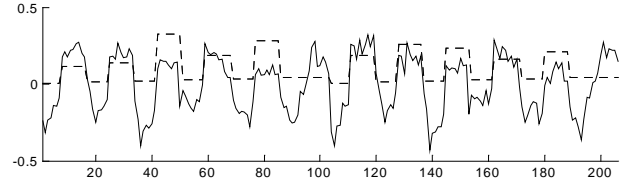


Figure 1. Time sequence extracted from 3-subject data that best correlates with the known external force stimulus (dashed line).

Our algorithm extracted this sequence without any prior knowledge of the experimental design. The images (columns of matrices \mathbf{B}_i) show how much each of the voxels follows this particular time pattern. Figure 2 shows three cross-sections of the resulting volume for the third subject associated with the time sequence shown in Figure 1. Dark areas are those whose activation levels over time follow the sequence in Figure 1.

To test the robustness of our algorithm to the misalignment intrinsic in multi-subject data sets we drastically spatially misaligned the third subject's data and repeated the analysis. The best extracted time sequence is almost identical to the one shown in Figure 1 – correlation coefficient 0.9982. Figure 3 shows the subject 3 volume from \mathbf{B}_3 associated with this time sequence. To compare the analysis results from the aligned and misaligned data sets we transformed the volume in Fig. 3 back to the original coordinates of Fig. 2. Figure 4 shows the scatter plot of voxel values of the misaligned-processed-realigned data versus the original data for subject 3. The scatter plot shows a strong image similarity with a correlation coefficient of 0.8854.

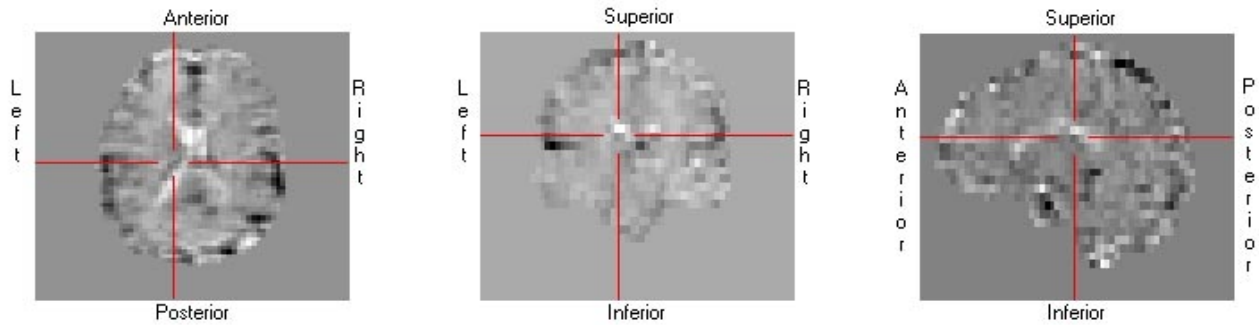


Figure 2. Subject 3 volume associated with the time sequence shown in Figure 1. Black areas are those whose activation levels over time follow the time sequence in Figure 1. The lines represent coronal, sagittal and transverse planes.

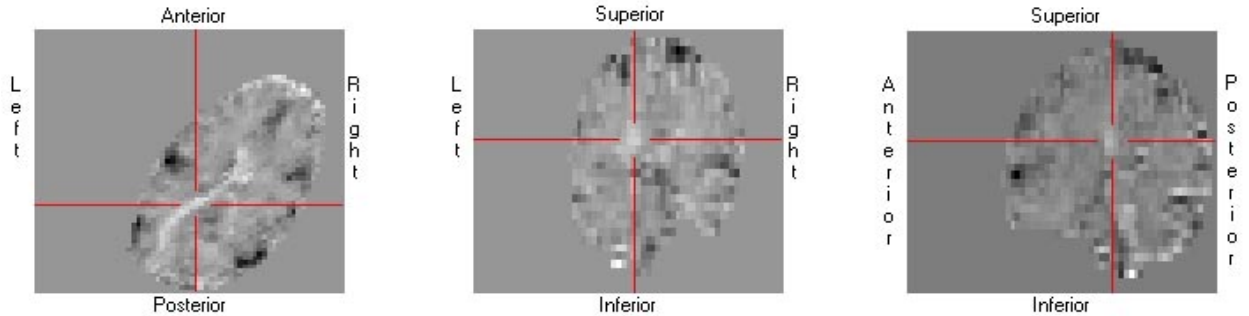


Figure 3. Subject 3 volume resulting from the analysis based on the spatially transformed data. Extracted time sequence was almost identical to the one extracted from aligned data.

6. CONCLUSIONS

In this paper we have extended the original blind source separation problem to include multiple, independent data sets, and presented an algorithm based on non-vanishing source autocorrelation functions for solving it. We assume that each data set is composed of a unique set of independent sources and another set of sources common to all data sets. We applied the algorithm to fMRI data sets from three subjects. Our algorithm readily extracted an “on-off” time sequence from the data that matched the experimental design. We also demonstrated that this algorithm is robust to subject misalignment, which might allow fMRI group analyses to be performed without intersubject registration to a common coordinate system.

Acknowledgements: This work was supported in part by NIH grants NS35273 and MH57180 and the Danish Research Councils through the THOR Center for Neuroinformatics.

7. REFERENCES

[1] Bell, A. and T. J. Sejnowski, “An Information-Maximization Approach to Blind Separation and Blind Deconvolution,” *Neural Computation*, vol. 7, pp. 1129-1159, 1995.
 [2] Cardoso, J.-F., A. Belouchrani, K. Abed-Meriam, and E. Moulines, “Blind Source Separation Technique Using Second-Order Statistics,” *IEEE Transactions on Signal Processing*, vol. 45, pp. 434-444, 1997.
 [3] Comon, P., “Independent Component Analysis: A New Concept ?,” *Signal Processing*, vol. 36, pp. 287-314, 1994.
 [4] LaConte, S., J. Anderson, S. Muley, J. Ashe, S. Fruitiger, K. Rehm, L. K. Hansen, E. Yacoub, X. Hu, D. Rottenberg, and

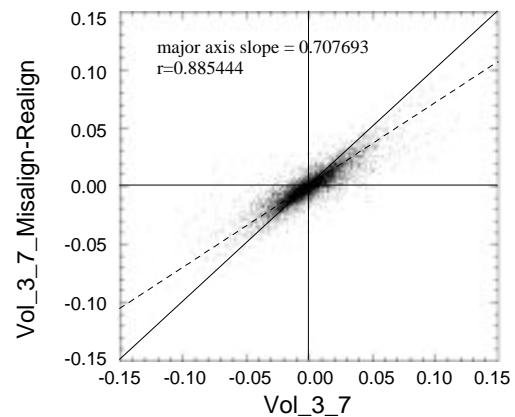


Figure 4. Scatter plot of voxel values of misaligned-processed-realigned data versus aligned-processed data for subject 3.

S. Strother, “The Evaluation of Preprocessing Choices in Single-Subject BOLD-fMRI Using NPAIRS Performance Metrics,” *Neuroimage*, (submitted), 2002.
 [5] Larsen, J., L. K. Hansen, and T. Kolenda, “On Independent Component Analysis for Multimedia Signals,” in *Multimedia Image and Video Processing*, G. L., K. S.Y., and J. Larsen, Eds.: CRC Press, 2000, pp. 175-200.
 [6] Molgedey, L. and H. Schuster, “Separation of Independent Sources Using Time-Delayed Correlations,” *Physical Review Letters*, vol. 72, pp. 3634-3637, 1994.
 [7] Petersen, K. S., L. K. Hansen, T. Kolenda, E. Rostrup, and S. C. Strother, “On the Independent Components of Functional Neuroimages,” 2000.
 [8] Proakis G. John and M. G. Dimitris, *Digital Signal Processing: Principles, Algorithms and Applications*, Third ed. Upper Saddle River, New Jersey: Prentice-Hall, 1996.