Effect of Spatial Alignment Transformations in PCA and ICA of Functional Neuroimages

Ana S. Lukic, *Member, IEEE,* Miles N. Wernick,* *Senior Member, IEEE,*
Yongyi Yang, *Senior Member, IEEE,* Lars Kai Hansen, *Member, IEEE,* Konstantinos Arfanakis, and
Stephen C. Strother, *Member, IEEE*

Abstract—It has been previously observed that spatial independent component analysis (ICA), if applied to data pooled in a particular way, may lessen the need for spatial alignment of scans in a functional neuroimaging study. In this paper we seek to determine analytically the conditions under which this observation is true, not only for spatial ICA, but also for temporal ICA and for principal component analysis (PCA). In each case we find conditions that the spatial alignment operator must satisfy to ensure invariance of the results. We illustrate our findings using functional magnetic-resonance imaging (fMRI) data. Our analysis is applicable to both inter-subject and intra-subject spatial normalization.

Index Terms—fMRI, Image Registration, Independent Component Analysis, Neuroimaging.

I. INTRODUCTION

IN functional neuroimaging it is standard practice to use spatial transformations to bring all the images into approximate spatial correspondence prior to their analysis [1]. This process, sometimes known as *spatial normalization*, is usually a necessary pre-processing step, because most analytical techniques assume that corresponding voxels in different images refer to the same location within the brain. Considerable attention is paid to minimizing inherent errors in the anatomical normalization process, and one can never say with certainty that intersubject normalization has been entirely

successful. Therefore, it may be desirable in certain types of studies to do away with spatial normalization. In this paper, we consider some situations in which this might be possible.

Svensén, *et al.*, [2] have mentioned that spatial independent component analysis (ICA) does not require spatial alignment, but they did not study the issue. In [3], we observed that temporal ICA may be invariant to spatial alignment operations, but we did not investigate the question either. There have been many papers based on different ways of forming the data matrix in ICA and PCA (e.g.,[3]-[7]), some of which may obviate the need for spatial alignment. However, to our knowledge, the specific conditions under which spatial ICA, temporal ICA, and principal component analysis (PCA) may be invariant to spatial alignment have not been studied.

In this paper, we investigate the conditions under which spatial normalization may be unnecessary when PCA or ICA is used to analyze the data. We show that, under some conditions, the temporal patterns obtained by PCA and ICA are invariant to spatial alignment transformations, and that these transformations have no effect on the spatial patterns except to align them. We present experimental data supporting our mathematical conclusions for one kind of PCA and one specific ICA algorithm.

In this paper, we will focus on functional magnetic-resonance imaging (fMRI) data; however the concepts apply equally to other functional-imaging modalities in which images are acquired in a time series, such as dynamic positron emission tomography (PET) or magnetoencephalography (MEG). We anticipate that the methods studied in this paper may be particularly appropriate for MEG, where spatial normalization prior to analysis is sometimes difficult.

The remainder of the paper is organized as follows. In the next section we discuss issues of data pooling in the context of PCA and ICA. In Sect. III, we discuss a matrix representation of the spatial-normalization transformation and sufficient conditions for invariance of the analysis. In Sect. IV, we examine these conditions for some basic spatial transformations. Comparisons of experimental PCA and ICA results with and without alignment are presented in Sect. V. Conclusions are given in Sect. VI.

Asterisk denotes corresponding author.

A. S. Lukic was with the Department of Biomedical Engineering, Illinois Institute of Technology, Chicago, IL, 60616, USA. She is now with Predictek, Inc, Chicago, IL 60616.

Y. Yang is with the Department of Electrical and Computer Engineering and Department of Biomedical Engineering, Illinois Institute of Technology, Chicago, IL, 60616, USA

L. K. Hansen is with the Informatics and Mathematical Modelling, Technical University of Denmark DK-2800 Lyngby, Denmark.

K. Arfanakis is with the Department of Biomedical Engineering, Illinois Institute of Technology, Chicago, IL, 60616, USA

S. C. Strother is with the Rotman Research Institute, Baycrest and University of Toronto, Toronto, ON, M6A-2E1, Canada.

*M. N. Wernick is with the Department of Electrical and Computer Engineering and Department of Biomedical Engineering, Illinois Institute of Technology, 3301 South Dearborn Street, Chicago, IL, 60616, USA (e-mail: wernick@iit.edu, phone: 312-567-8818, fax: 312-567-8976).

II. DATA MATRIX AND THE SPATIAL NORMALIZATION TRANSFORMATION

Let us assume that a sequence of N images is acquired for each of two subjects or two runs of the same subject. (We confine our attention to this case for notational simplicity; however, our analysis can readily be extended to more than two sets of images.) Let the two images at discrete time i be denoted by $\mathbf{x}_i$ and $\mathbf{y}_i$, $i = 1, \ldots, N$, which are column vectors formed by lexicographic ordering of the voxel values. Let $\mathbf{X}$ and $\mathbf{Y}$ denote data matrices formed from these image vectors as follows: $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_N \end{bmatrix}$ and $\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_N \end{bmatrix}$.

To analyze these data sets simultaneously, it is common in neuroimaging to pool the data into a single data matrix as follows [7]-[10]:

$$\begin{aligned} \mathbf{Z} &= \begin{bmatrix} \mathbf{X} & \mathbf{Y} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_N & \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_N \end{bmatrix} \end{aligned} \quad (1)$$

In this case, spatial normalization is necessary because all the values within a given row of $\mathbf{Z}$ are assumed to refer to the same spatial location.

In this paper we will analytically study alternative ways of forming the data matrix $\mathbf{Z}$ [2], which we will denote by $\mathbf{Z}_1$ and $\mathbf{Z}_2$, so that either:

$$\mathbf{Z} = \mathbf{Z}_1 = \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}, \quad (2)$$

or

$$\mathbf{Z} = \mathbf{Z}_2 = \begin{bmatrix} \mathbf{X}^T & \mathbf{Y}^T \end{bmatrix}. \quad (3)$$

It is important to note that these two representations are simply transposes of one another, i.e., $\mathbf{Z}_1 = \mathbf{Z}_2^T$. While these two representations are equivalent when PCA is applied; in ICA the data matrix in (2) yields temporal ICA, while the data matrix in (3) produces spatial ICA.

In this paper, we will consider spatial-normalization transformations that can be expressed as a linear transformation $\mathbf{R}$ of the voxel values, typically having the form

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_X & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_Y \end{pmatrix}, \quad (4)$$

in which $\mathbf{R}_X$ and $\mathbf{R}_Y$ are transformations that spatially transform the data in $\mathbf{X}$ and $\mathbf{Y}$, respectively, usually to a standard template.

It is important to note that we are not assuming that the transformation operator $\mathbf{R}$ is linear in terms of the spatial coordinates in the image domain; it only must be linear in terms of the voxel values. Thus, so-called "nonlinear warping" methods are not excluded by this assumption. However, our assumption does require that linear interpolation be used when implementing these warping transformations.

When using the data matrix in (2), the operation of spatial alignment is denoted by:

$$\mathbf{Z}_A = \mathbf{R}\mathbf{Z} ; \quad (5)$$

whereas, when using the data matrix in (3), spatial alignment is represented as:

$$\mathbf{Z}_A = \mathbf{Z}^T \mathbf{R}^T . \quad (6)$$

Here, and throughout the paper, variables with subscript A signify quantities associated with post-aligned data.

We point out that the same setup can be used to analyze the effect of temporal transformations (e.g. phase shift) on data pooled along the temporal direction. In this case matrix $\mathbf{R}$ represents temporal transformations and the transformed data are written as $\mathbf{Z}_A = \mathbf{Z}\mathbf{R}$. Applying ICA on $\mathbf{Z}_A = \mathbf{Z}\mathbf{R}$ yields temporal independent components and, as above, ICA of $\mathbf{Z}_A = \mathbf{R}^T \mathbf{Z}^T$ yields spatial independent components.

In the next section, we will show that, using the data matrix formulations in (2) and (3), PCA and ICA are invariant to certain kinds of spatial transformations. Specifically, we will show that, in these situations, the temporal patterns are unaffected by alignment procedures. We also show that, while these transformations serve to align the spatial patterns of the subjects, they do not otherwise change the essential nature of the spatial patterns.

III. INVARIANCE TO ALIGNMENT TRANSFORMATIONS

We will consider analysis methods that decompose the data matrix into a spatial matrix $\mathbf{G}$ and a temporal matrix $\mathbf{T}$, i.e., $\mathbf{Z} = \mathbf{GT}$. Among the best known decompositions of this type are PCA [11] and ICA [12].

A. Principal Component Analysis (PCA)

The two data representations defined in (2) and (3) are simply transposes of one another; thus, in regards to PCA, the difference between the two representations is merely one of notation. Therefore, in the following analysis of the PCA method we will only consider the arrangement in (2).

Let $\mathbf{Z} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$ denote a singular value decomposition (SVD) of $\mathbf{Z}$. Given that, in neuroimaging, $\mathbf{Z}$ typically has many more rows than columns (more voxels than time samples), we will consider the reduced SVD of $\mathbf{Z}$, in which

the diagonal matrix $\mathbf{\Lambda}$ contains only the nonzero singular values of $\mathbf{Z}$, and the matrices $\mathbf{U}$ and $\mathbf{V}$ are formed from the corresponding left and right singular vectors, respectively. Then the rows of $\mathbf{T}$ are formed from the rows of the matrix $\mathbf{V}^T$, and the spatial patterns are the columns of the matrix

$$\mathbf{G} = \mathbf{U}\mathbf{\Lambda} = \begin{pmatrix} \mathbf{U}_X \\ \mathbf{U}_Y \end{pmatrix} \mathbf{\Lambda} . \tag{7}$$

Note that the spatial patterns obtained via PCA are always mutually orthogonal, i.e., $\mathbf{G}^T\mathbf{G} = \mathbf{\Lambda}^2$ is a diagonal matrix.

With the data arranged as in (2), the temporal patterns obtained by PCA are unaffected by certain spatial transformations, and the resulting spatial patterns are simply aligned versions of the patterns that would be obtained without pre-alignment. This is easily shown as follows. Applying the spatial transformation $\mathbf{R}$ to the data, we obtain:

$$\mathbf{Z}_A = \mathbf{R}\mathbf{Z} = \mathbf{R}\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T = \mathbf{G}_A \mathbf{V}^T , \tag{8}$$

where

$$\mathbf{G}_A = \mathbf{R}\mathbf{U}\mathbf{\Lambda} . \tag{9}$$

Equation (8) shows the representation of the aligned data in which temporal patterns are the same as those obtained from the unaligned data by the PCA. Equations (7) and (9) show that the spatial patterns in (8) are simply aligned versions of the spatial patterns in $\mathbf{G}$, i.e.,

$$\mathbf{G}_A = \mathbf{R}\mathbf{G} . \tag{10}$$

Now let us denote the SVD of $\mathbf{Z}_A$ as $\mathbf{Z}_A = \mathbf{U}_A\mathbf{\Lambda}_A\mathbf{V}_A^T$. From (8) it can be seen that we will have $\mathbf{V}_A = \mathbf{V}$ provided that the following condition is met:

$$\left(\mathbf{R}\mathbf{U}\right)^T \left(\mathbf{R}\mathbf{U}\right) = \mathbf{U}^T\mathbf{R}^T\mathbf{R}\mathbf{U} = \lambda\mathbf{I} . \tag{11}$$

where $\lambda > 0$ is an arbitrary constant. Moreover, in such a case we have $\mathbf{\Lambda}_A = \sqrt{\lambda}\mathbf{\Lambda}$ and $\mathbf{U}_A = \mathbf{R}\mathbf{U}/\sqrt{\lambda}$.

The condition in (11) requires that $\mathbf{R}$ be a unitary transformation (up to an arbitrary scale factor) on the subspace spanned by the left singular vectors of $\mathbf{Z}$ (i.e., the range space of $\mathbf{U}$). Of course, this condition is easily met if

$$\mathbf{R}^T\mathbf{R} = \lambda\mathbf{I} . \tag{12}$$

However, it should be noted that the condition in (12) is much more restrictive than that in (11) since, in neuroimaging,

the dimension of $\mathbf{U}$ is much lower than the number of rows of $\mathbf{Z}$. Even so, we demonstrate that the condition in (12) is satisfied by some general geometric transformations.

To summarize, PCA is essentially unaffected by spatial alignment so long as the alignment transformation (expressed in terms of the voxel values) can be represented by a matrix $\mathbf{R}$ obeying property (11). In Sect. IV, we will show that important types of transformations either exactly or approximately satisfy this requirement. Next we consider the effect of spatial normalization on ICA.

B. Independent Component Analysis (ICA)

ICA refers broadly to a collection of techniques that aim to decompose data into components that are statistically independent. Because statistical independence of a random sequence is a strong statement, one usually cannot fully specify or guarantee independence in practice. Instead, the usual approach is to seek components that optimize some manifestation of independence using measures such as lagged correlation [13], mutual information [12], negentropy [12], or kurtosis [12].

In general, the ICA data model is

$$\mathbf{Z} = \mathbf{A}\mathbf{S} , \tag{13}$$

where $\mathbf{Z}$ is the data matrix, $\mathbf{\Lambda}$ is the so-called *mixing matrix*, and $\mathbf{S}$ is a matrix of statistically independent *sources*. ICA is often referred to as *blind source separation*, because the mixing matrix $\mathbf{A}$ and the source matrix $\mathbf{S}$ are assumed to be unknown. In spatial ICA, the sources are spatial patterns; in temporal ICA, they are temporal patterns.

Because the two data representations in (2) and (3) lead to temporal and spatial ICA respectively, we analyze them separately.

C. Spatial ICA

Using the data matrix in (3) in conjunction with (13) yields a spatial ICA, i.e., an analysis in which it is assumed that brain activation is driven by a set of statistically independent spatial patterns, weighted by time factors contained in the mixing matrix $\mathbf{A}$.

In this formulation, if a spatial transformation $\mathbf{R}$ is applied to the image data, then the ICA model in (13) becomes:

$$\mathbf{Z}_A = \mathbf{Z}\mathbf{R}^T = \mathbf{A}\mathbf{S}\mathbf{R}^T = \mathbf{A}\left(\mathbf{S}\mathbf{R}^T\right). \tag{14}$$

Now we ask the following: Is $\left(\mathbf{S}\mathbf{R}^T\right)$ a valid matrix of independent spatial patterns? The answer is affirmative. That is, we have $\mathbf{S}_A = \mathbf{S}\mathbf{R}^T$, and the mixing matrix $\mathbf{A}$ is unchanged. To see this, write $\mathbf{S}$ as

$$\mathbf{S} = [\mathbf{s}_1^T \quad \mathbf{s}_2^T \quad \cdots \quad \mathbf{s}_m^T]^T , \tag{15}$$

where $\mathbf{s}_j$, $j = 1, \ldots, m$, are row vectors denoting the independent spatial components of $\mathbf{Z}$. Then,

$$\mathbf{SR}^T = \begin{bmatrix} \mathbf{s}_1 \mathbf{R}^T \\ \mathbf{s}_2 \mathbf{R}^T \\ \vdots \\ \mathbf{s}_m \mathbf{R}^T \end{bmatrix}. \tag{16}$$

Clearly, the rows of this matrix remain independent provided that the rows of $\mathbf{S}$ are independent. However, note that in order to obtain $\mathbf{S}$ from $\mathbf{S}_A$, the matrix $\mathbf{R}$ must be invertible.

Thus, in general, invariance of spatial ICA results requires only the invertibility of $\mathbf{R}$. However, we will see later, the requirements on $\mathbf{R}$ may be more strict for any given computational algorithm for performing ICA.

D. Temporal ICA

The data matrix in (2) yields a temporal ICA, in which brain activation is modeled as a mixture of independent temporal components, each weighted by a corresponding spatial pattern contained in the mixing matrix $\mathbf{A}$.

Now let us consider the effect of spatial normalization on temporal ICA. If a spatial transformation $\mathbf{R}$ is applied to the image data, then the ICA model in (13) becomes:

$$\mathbf{Z}_A = \mathbf{RZ} = \mathbf{RAS}. \tag{17}$$

Thus, the original ICA model in (13) has been transformed to a new ICA model, which can be expressed as:

$$\mathbf{Z}_A = \mathbf{A}_A \mathbf{S}, \tag{18}$$

wherein $\mathbf{A}_A = \mathbf{RA}$ and the source matrix $\mathbf{S}$ is preserved. Since ICA imposes no further conditions on the matrices, (18) demonstrates that the spatial transformation $\mathbf{R}$ can be applied before or after the analysis with identical results, provided that $\mathbf{R}$ is invertible. Therefore, in principle, it appears that this kind of ICA is essentially invariant to spatial alignment transformations as well; however, as we will see next, this may not be the case for any given ICA algorithm.

E. Molgedey-Schuster ICA Algorithm

It is important to note that the goal of ICA (to find statistically independent components) is never precisely achieved in practice, because independence is a statement about probability density functions of all orders $n = 1, \ldots, \infty$, a condition that cannot be discerned by any realizable numerical algorithm. In practice, ICA algorithms merely seek to optimize some signature of independence; therefore, the peculiar characteristics of any given ICA algorithm may lead to results that are not invariant to the same broad class of spatial transformations. To investigate this possibility, we next consider one specific algorithm, known as the Molgedey-Schuster (M-S) method [13],[14]. We choose to study this particular algorithm because it is a simple non-iterative method, which makes our analysis tractable, and because it has been applied successfully in the past [14],[3],[15].

ICA algorithms use various manifestations of independence to find components that are purportedly independent. In the M-S method, sources are considered to be independent if their lagged cross correlation is zero. In concept, the M-S algorithm consists of two steps: a whitening followed by a rotation. In the whitening step, a PCA transformation is applied to the data. This yields components that are uncorrelated, but not necessarily independent. In the rotation step, the time-lagged covariance matrix of the basis vectors is diagonalized in an effort to obtain independent sources.

Now let us consider temporal ICA implemented using the M-S algorithm. In the Appendix, we show that the temporal M-S ICA algorithm is equivalent to computing the mixing matrix $\mathbf{A}$ by solving the following eigenvector equation:

$$\mathbf{C}_Z(\tau)\left(\mathbf{U}\mathbf{\Lambda}^{-2}\mathbf{U}^T\right)\mathbf{A} = \mathbf{A}\mathbf{C}_S(\tau), \tag{19}$$

in which $\mathbf{C}_Z(\tau)$ and $\mathbf{C}_S(\tau)$ are the time-lagged autocorrelation matrices of the data and the sources, respectively. It should be noted that $\mathbf{C}_S(\tau)$ is a diagonal matrix.

Let us suppose that we implement (19) using spatially aligned data $\mathbf{Z}_A$ in place of the original data $\mathbf{Z}$. In this case, the time-lagged covariance matrix of $\mathbf{Z}_A$ is related to that of $\mathbf{Z}$ by:

$$\mathbf{C}_{Z_A}(\tau) = \mathbf{R}\mathbf{C}_Z(\tau)\mathbf{R}^T. \tag{20}$$

Now let us write the reduced SVD of $\mathbf{Z}_A$ as $\mathbf{Z}_A = \mathbf{U}_A \mathbf{\Lambda}_A \mathbf{V}_A^T$. Then, the eigenvector equation for the mixing matrix becomes:

$$\mathbf{C}_{Z_A}(\tau)\left(\mathbf{U}_A \mathbf{\Lambda}_A^{-2}\mathbf{U}_A^T\right)\mathbf{A}_A = \mathbf{A}_A \mathbf{C}_A(\tau), \tag{21}$$

where $\mathbf{C}_A(\tau)$ is the time-lagged autocorrelation matrix of the (new) sources.

Now suppose that the transformation $\mathbf{R}$ satisfies the condition in (11). Then, we have $\mathbf{\Lambda}_A = \sqrt{\lambda}\mathbf{\Lambda}$ and $\mathbf{U}_A = \mathbf{RU}/\sqrt{\lambda}$. Substituting these relations into (20) and (21), we obtain:

$$\mathbf{R}\mathbf{C}_z\left(\tau\right)\mathbf{R}^T\left(\frac{1}{\sqrt{\lambda}}\mathbf{R}\mathbf{U}\frac{1}{\lambda}\mathbf{\Lambda}^{-2}\frac{1}{\sqrt{\lambda}}\mathbf{U}^T\mathbf{R}^T\right)\mathbf{A}_A =$$
$$= \mathbf{A}_A\mathbf{C}_A\left(\tau\right) \tag{22}$$

Upon further algebraic manipulation, we obtain

$$\mathbf{C}_z\left(\tau\right)\left(\mathbf{U}\mathbf{\Lambda}^{-2}\mathbf{U}^T\right)\mathbf{R}^{-1}\mathbf{A}_A = \mathbf{R}^{-1}\mathbf{A}_A\mathbf{C}_A\left(\tau\right). \tag{23}$$

Thus,

$$\mathbf{A}_A = \mathbf{R}\mathbf{A}. \tag{24}$$

Therefore, the mixing matrix computed from the aligned data is simply an aligned version of what would have been found from the original data (up to permutation and sign changes).

Note that, to obtain (23), it is assumed that the transformation $\mathbf{R}$ is invertible and satisfies the condition in (11).

Now we consider the effect of spatial normalization on the temporal patterns (sources). As we discuss in the Appendix, the Molgedey-Schuster method determines the sources as $\mathbf{S} = \mathbf{Q}\mathbf{V}^T$, where $\mathbf{Q}$ is the eigenvector matrix of the time-lagged covariance matrix of $\mathbf{V}$. We have already shown in Sect. III.A that $\mathbf{V}$ is invariant to the spatial transformation $\mathbf{R}$ provided that it satisfies the condition in (11), therefore $\mathbf{V}_A = \mathbf{V}$. Since $\mathbf{Q}$ is derived solely and uniquely from $\mathbf{V}$, it immediately follows that $\mathbf{Q}_A = \mathbf{Q}$. Therefore, the sources derived from the aligned data are identical to those derived from the original data, i.e., $\mathbf{S}_A = \mathbf{Q}_A\mathbf{V}_A^T = \mathbf{Q}\mathbf{V}^T = \mathbf{S}$.

Thus, we have demonstrated that the temporal M-S ICA algorithm is invariant to any spatial transformation $\mathbf{R}$ that satisfies the condition in (11), which is more restrictive than what we might expect in principle for ICA, as shown in Sect. III.D. However, it should be noted that if the number of time samples is greater than or equal to the number of voxels, then invariance of the M-S algorithm only requires that $\mathbf{R}$ be invertible, which is the same condition we found for ICA in general. However, this less-restrictive case is not often of interest in neuroimaging.

IV. PROPERTIES OF THE ALIGNMENT TRANSFORMATION

In the previous section we considered conditions under which PCA and ICA are essentially unaffected by spatial alignment transformations. Next we examine whether these conditions are met by actual transformations.

For notational convenience, we will consider two-dimensional images; however, the results can be readily extended to three-dimensional images. Furthermore, we will assume that the image functions of interest are elements of L^2

(the space of square-integrable functions). Consider an image function $f\left(x,y\right)$ of continuous spatial variables x and y, and let $\tilde{\mathbf{R}}$ denote the operator on $f\left(x,y\right)$ corresponding to either spatial transformation $\mathbf{R}_X$ or $\mathbf{R}_Y$ in (2). Let $f_A\left(x,y\right)$ denote $f\left(x,y\right)$ after application of $\tilde{\mathbf{R}}$, i.e., $f_A\left(x,y\right) = \tilde{\mathbf{R}}\left[f\left(x,y\right)\right]$. For clarity, we emphasize that $\tilde{\mathbf{R}}$ is a linear operator defined on image functions in L^2; in the case of discrete images, this operator is represented in matrix form as in our discussions in the previous section.

A. Affine Transforms

In its most general form, an affine transform can be represented as

$$\begin{pmatrix}x'\\y'\end{pmatrix} = \begin{pmatrix}a_1 & a_2\\a_3 & a_4\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} + \begin{pmatrix}b_1\\b_2\end{pmatrix}, \tag{25}$$

where $\left(x,y\right)$ and $\left(x',y'\right)$ are spatial coordinates before and after the transformation, respectively, and a_i and b_j, $i = 1,2,3,4$, $j = 1,2$, are transform parameters. Recall that the affine transform describes combinations of anisotropic scaling, shear, rotations, and translations.

For any two images $f\left(x,y\right), g\left(x,y\right) \in L^2$, it can be shown, subject to the transformation in (25), that the following identity holds:

$$\begin{aligned}\langle f_A(x,y), g_A(x,y)\rangle &= \iint_D f_A(x,y)g_A(x,y)dxdy\\&= \left|a_1a_4 - a_2a_3\right|\langle f(x,y), g(x,y)\rangle\end{aligned}, \tag{26}$$

where $\langle\cdot,\cdot\rangle$ denotes an inner product. Consequently, for an affine transformation $\tilde{\mathbf{R}}$, we have

$$\tilde{\mathbf{R}}^T\tilde{\mathbf{R}} = \left|a_1a_4 - a_2a_3\right|\mathbf{I}, \tag{27}$$

where $\tilde{\mathbf{R}}^T$ denotes the adjoint operator of $\tilde{\mathbf{R}}$, and the scalar factor $\left|a_1a_4 - a_2a_3\right|$ corresponds to the area ratio associated with the transform. For transformations involving only rotation and/or translation this factor is equal to 1 since these operations do not change the size of the brain region corresponding to each pixel. If the image is scaled by factors s_x and s_y in x and y directions respectively, the factor is equal to $s_x s_y$. Therefore, the property in (12) holds for any combination of global anisotropic scaling, shear, rotation, and translation.

B. Neuroimaging Alignment Transformations

Of course, neuroimaging alignment transformations are not restricted to affine transformations [16],[9]. Nevertheless, it is always possible to approximate a given spatial transformation using piecewise affine transforms over small regions. Indeed, in the Talairaich coordinate system, the brain can be divided into regions, to which affine transforms can be fit separately [17].

Of course, an issue that arises from the use of piecewise affine approximation is that the factor $|a_1 a_4 - a_2 a_3|$ in the identity in (20) now may differ for the different pieces. However, in practice we expect that while such variations will exist, their extent will be limited and may cause a negligible effect on the analysis, as will be seen from our experiments.

A similar issue is that so far we have limited the spatial transformation to a single image. However, in (2) the transformation $\mathbf{R}$ involves two images, each of which is associated with its own spatial transform. Nevertheless, one can see that $\mathbf{R}^T \mathbf{R} = \lambda \mathbf{I}$ also holds approximately in such a case, provided that the factor $|a_1 a_4 - a_2 a_3|$ is not significantly different for the two spatial transforms.

C. Discrete Images

In the above derivations we have treated images as functions of continuous spatial variables. However, in practice we work with their discrete samples. One may wonder if the properties of the spatial transformation $\mathbf{R}$ would hold equally for discrete samples. The answer is affirmative, which we justify below.

Consider an image function $f(x, y)$. Let $f[m, n]$ denote its sampled version, i.e.,

$$f[m, n] = f(m\Delta x, n\Delta y),\qquad(28)$$

where $\Delta x, \Delta y$ are the sampling intervals in the x and y directions, respectively. Similarly, let $g[m, n]$ denote the sampled version of another image function $g(x, y) \in L^2$. Now let us assume that the images are properly sampled so that no aliasing occurs. Then from sampling theory the following identity can be derived:

$$\langle f[m, n], g[m, n] \rangle = \sum_{m,n} f[m, n] g[m, n]$$
$$= \frac{1}{\Delta x \Delta y} \iint_D f(x, y) g(x, y) dx dy \qquad(29)$$
$$= \frac{1}{\Delta x \Delta y} \langle f(x, y), g(x, y) \rangle.$$

Based on (29) we can see that the identity (27) is equally applicable for discrete images with the operator $\mathbf{R}$ now represented as a matrix.

We point out that the identity in (29) is based on the assumption that both image functions $f(x, y)$ and $g(x, y)$ are sampled without aliasing (i.e., without violating the Nyquist condition). Of course, this condition could potentially be violated when an image is scaled by an exceedingly large factor during alignment. However, in practice the inter-subject variation doesn't seem to have much impact, as illustrated in our experiments (Experiment 1). In addition, non-ideal interpolation is often used in practice because of its simplicity (as in our experiments), but our experimental results seem to suggest that it has little effect. Finally, we point out all the derivation results still hold for images square or not.

V. EXPERIMENTAL RESULTS

In this section we investigate the effect of spatial alignment transformations by comparing the results of analyses based on aligned and misaligned fMRI images, using PCA and temporal ICA. In the first experiment we used two runs of an fMRI study for the same subject, then deliberately applied random spatial transformations to one of the runs to create severely distorted and misaligned data. In this way we eliminated confounding effects of intersubject functional variations and were able to study the effects of extreme misalignments. In the second experiment we used the fMRI data from two subjects and compared the results of PCA and ICA analyses of the original (misaligned) and aligned data.

A. Experiment 1: two runs of the fMRI data for the same subject

The images were obtained while a right-handed volunteer performed two runs of a static force task, alternating six rest and five force periods per run (44 s/period; 200, 400, 600, 800, 1000g force levels between thumb and forefinger, pseudo-randomized across force periods and maintained with visual feedback) [18]. Images were corrected for within-subject motion and spatially aligned as described in [10], and the data matrix $\mathbf{Z}$ was doubly centered (to have zero mean).

We applied PCA and temporal M-S ICA to the two runs of fMRI data in two ways: first with the images spatially aligned, then again after the images were deliberately transformed spatially to simulate intersubject misalignment. Intersubject misalignment was modeled by a rotation by random angle θ, translation by a random vector $(\Delta x, \Delta y)$ and independent scale parameters a and b in the horizontal and vertical directions respectively, which permitted fairly extreme changes in the aspect ratio of the brain between the two fMRI runs. The transformation parameters were drawn from independent Gaussian distributions. The statistics (mean, standard deviation) of $\theta, \Delta x, \Delta y, a$ and b were, respectively, $(0°, 30°)$, (0 pixels, 5 pixels), (0 pixels, 5 pixels), and (1, 0.2) for both a and b. Example scans from 10 realizations of this random spatial transformation are shown in Fig. 1.
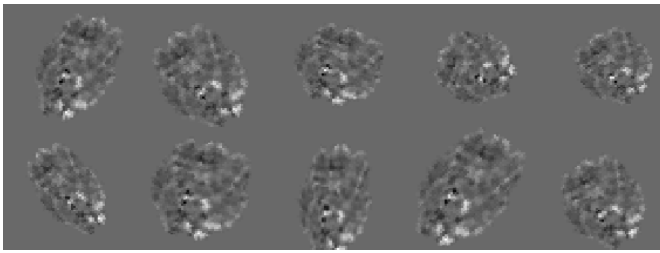
Fig. 1. Example scans from one run of fMRI data after random spatial transformations were introduced to simulate misalignments and intersubject anatomical variations, consisting of random changes of aspect ratio, rotations and translations.

Next, we applied PCA and Molgedey-Schuster ICA to the aligned and misaligned data sets, then compared the results using correlation and visual inspection. A quantitative comparison of the temporal patterns computed from aligned and misaligned data was made by computing the Pearson product-moment correlation between the patterns for 10 realizations of the misalignment transformation. The results of

TABLE I
CORRELATION BETWEEN PATTERNS COMPUTED WITH AND WITHOUT
SPATIAL ALIGNMENT

Method	Component	Temporal patterns	Spatial patterns
PCA	1	0.975 ± 0.016	0.983 ± 0.012
	2	0.995 ± 0.004	0.992 ± 0.006
	3	0.999 ± 0.001	0.998 ± 0.002
ICA	1	0.983 ± 0.012	0.971 ± 0.023
	2	0.992 ± 0.006	0.948 ± 0.005
	3	0.998 ± 0.002	0.999 ± 0.000

these comparisons are given in Table I. As expected, all the correlations are nearly one.

The approximate invariance of the results can also be seen from the example temporal patterns shown in Fig. 2 (left), in which the solid curves were computed from spatially aligned data and the dashed curves were obtained from misaligned data. The curves are virtually identical. In Fig. 2 (right) we show the spatial patterns obtained in this example. The spatial patterns obtained from the original data analysis were spatially transformed for ease of comparison.
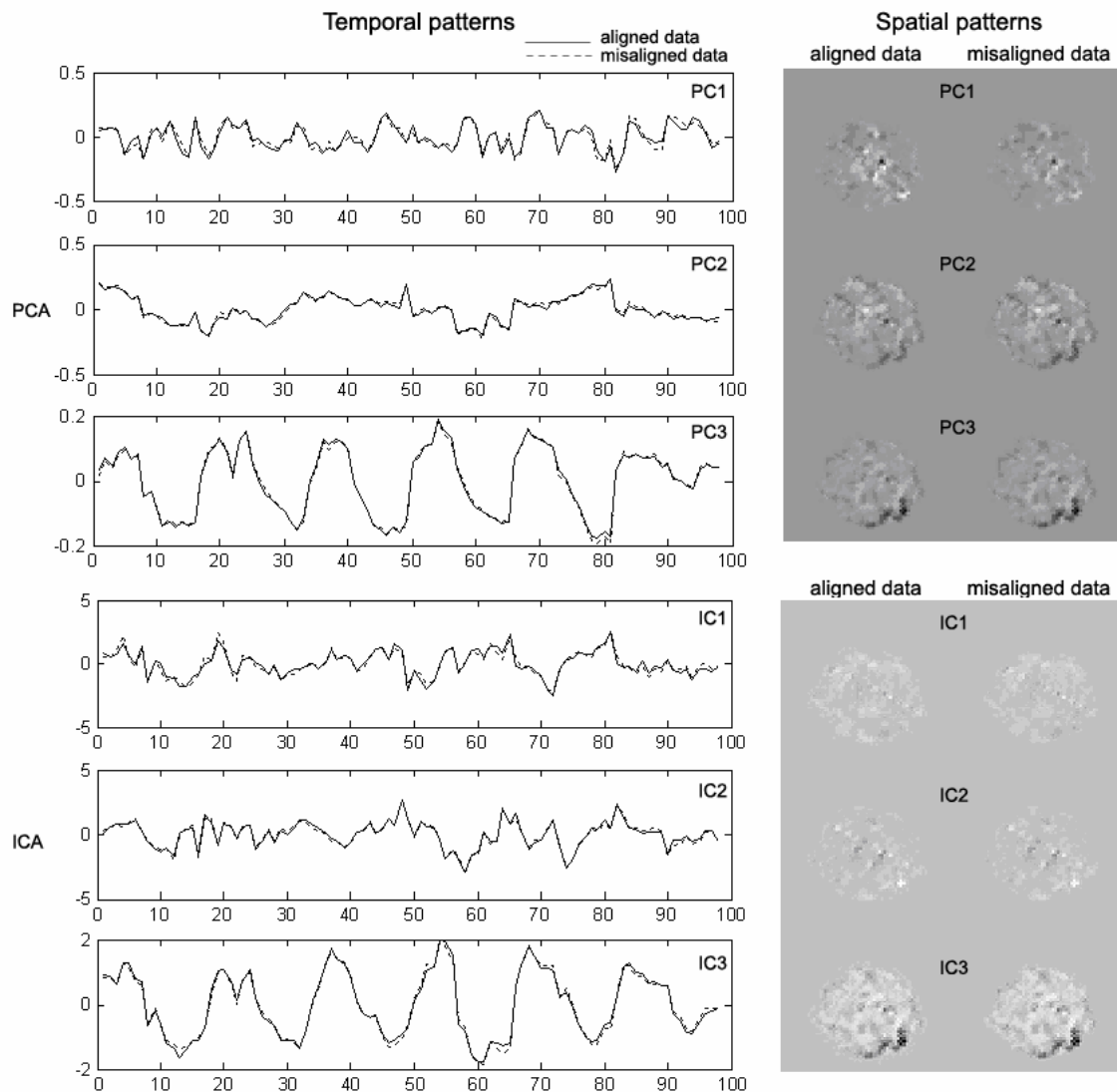

Fig. 2. First three temporal and spatial patterns computed with and without pre-alignment of the data by PCA and ICA. In this case, misalignments consisted of random changes in aspect ratio (anisotropic scaling), rotations, and translations. Solid curves were computed from aligned data, dashed curves from misaligned data. Spatial patterns were spatially transformed for ease of comparison. Spatial alignment appears to have virtually no effect on the results, an observation that is quantified by the correlation values in Table I.
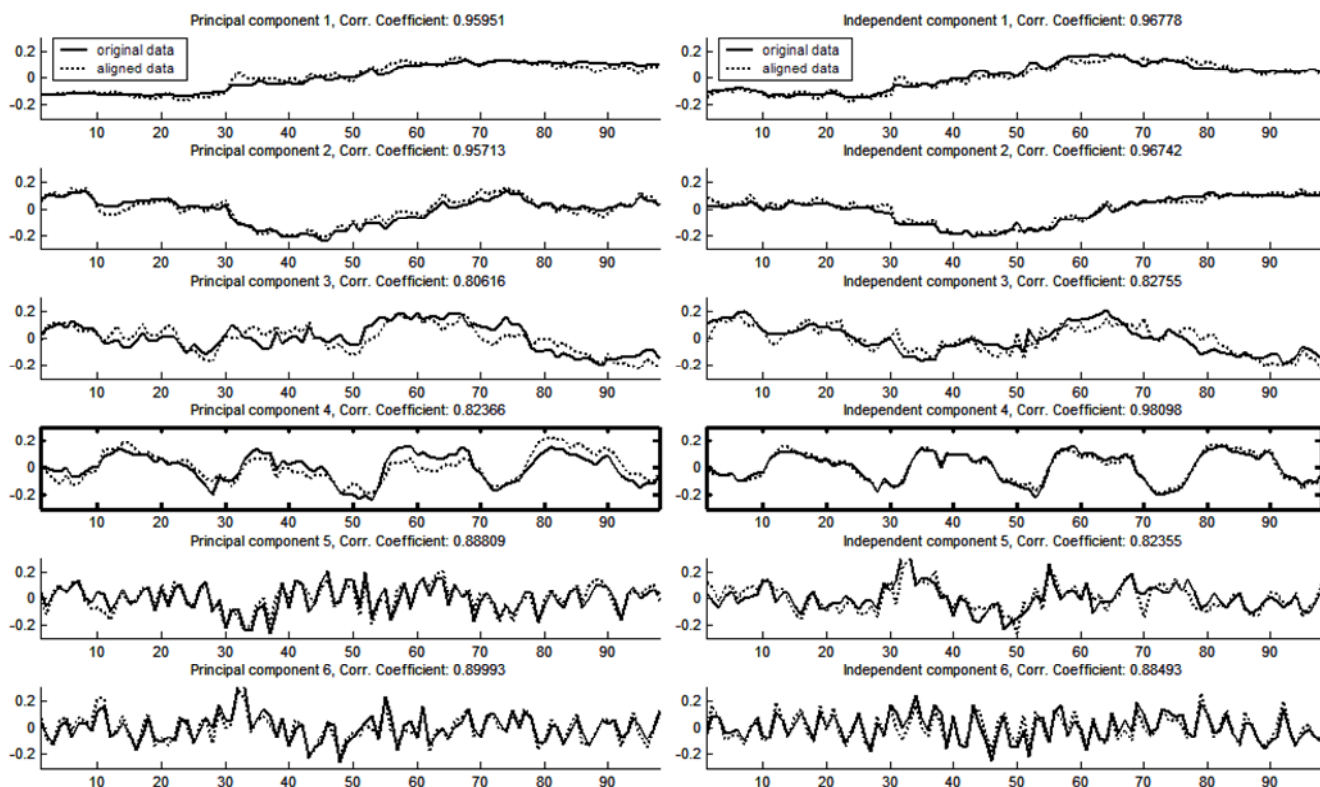
Fig. 3. The results of PCA and ICA analyses of the original and spatially normalized (aligned) data. The first six principal components account for more then 60% of the total variance in the data. Principal component 4 shows the highest correlation with the experimental paradigm.

B. *Experiment 2: fMRI data for two subjects*

In the next experiment we compare the results of PCA and ICA analyses on the fMRI data of two subjects with and without spatial normalization. This dataset was obtained while the subjects performed a visual and auditory task. During the study four task periods of 32 seconds were alternated with five rest periods of 20 seconds for the total duration of 228 seconds during which a total of 118 images were obtained. During the first task period, 8 images of the Wada objects were back projected onto a screen at the rate of one every four seconds. Simultaneously, the subject heard the name of the object repeated via earphones. During the second and third task periods, the subject was shown a series of eight images consisting of some that he/she has seen in the first task period (targets) and some unfamiliar ones (foils). The subject also heard the names of the objects. During the fourth task period, the subject was shown a series of eight unfamiliar objects. The subject was instructed to squeeze a pneumatic bulb when he/she recognized an object that was presented during the first task period. During the rest periods, the subjects were instructed to fixate their eyes on a white cross hair pattern. The imaging parameters were: TR/TE = 2000ms/40ms, FOV= 24cm, 7mm slice thickness, and 1-2mm interslice gap resulting in an 64 x 64 pixel image matrix and 24 coronal slices and voxel size of 3.75mm x 3.75mm x 7mm.

Prior to the analysis the data was pre-processed in the following way. The first four scans as well as one transitional scan from the beginning and end of each task/rest period were

TABLE II
CORRELATION BETWEEN PATTERNS COMPUTED WITH AND WITHOUT SPATIAL ALIGNMENT FROM THE AUDIO-VISUAL TASK DATASET

Method	Component	Correlation Coefficient	Component	Correlation Coefficient
PCA	1	0.95951	4	0.82366
	2	0.95713	5	0.88809
	3	0.80616	6	0.89993
ICA	1	0.96778	4	0.98098
	2	0.96742	5	0.82355
	3	0.82755	6	0.88493

discarded reducing the number of volumes per subject to 98. Each voxel time course was then made zero mean and unit variance. Finally, a 3D Gaussian smoothing kernel with the full with at half maximum (FWHM) of 8mm in each direction was applied to each volume. Data was then spatially normalized using the SPM5 software to the EPI template resulting in a 99 x 89x 115 voxel volumes and voxel size of 2mm x 2mm x 2mm.

In Fig. 3 we compare the first six temporal patterns obtained by PCA and ICA from original and spatially aligned data. Solid curves correspond to temporal patterns obtained from the original data while dotted curves correspond to temporal patterns obtained from the aligned data. These first six principal components account for more then 60% of the total variance in the data. The correlation coefficients between the patterns obtained with and without spatial normalization are given in Table II. As expected the correlation coefficients are very high although not as high as in the previous experiment where the two-dimensional data of the same
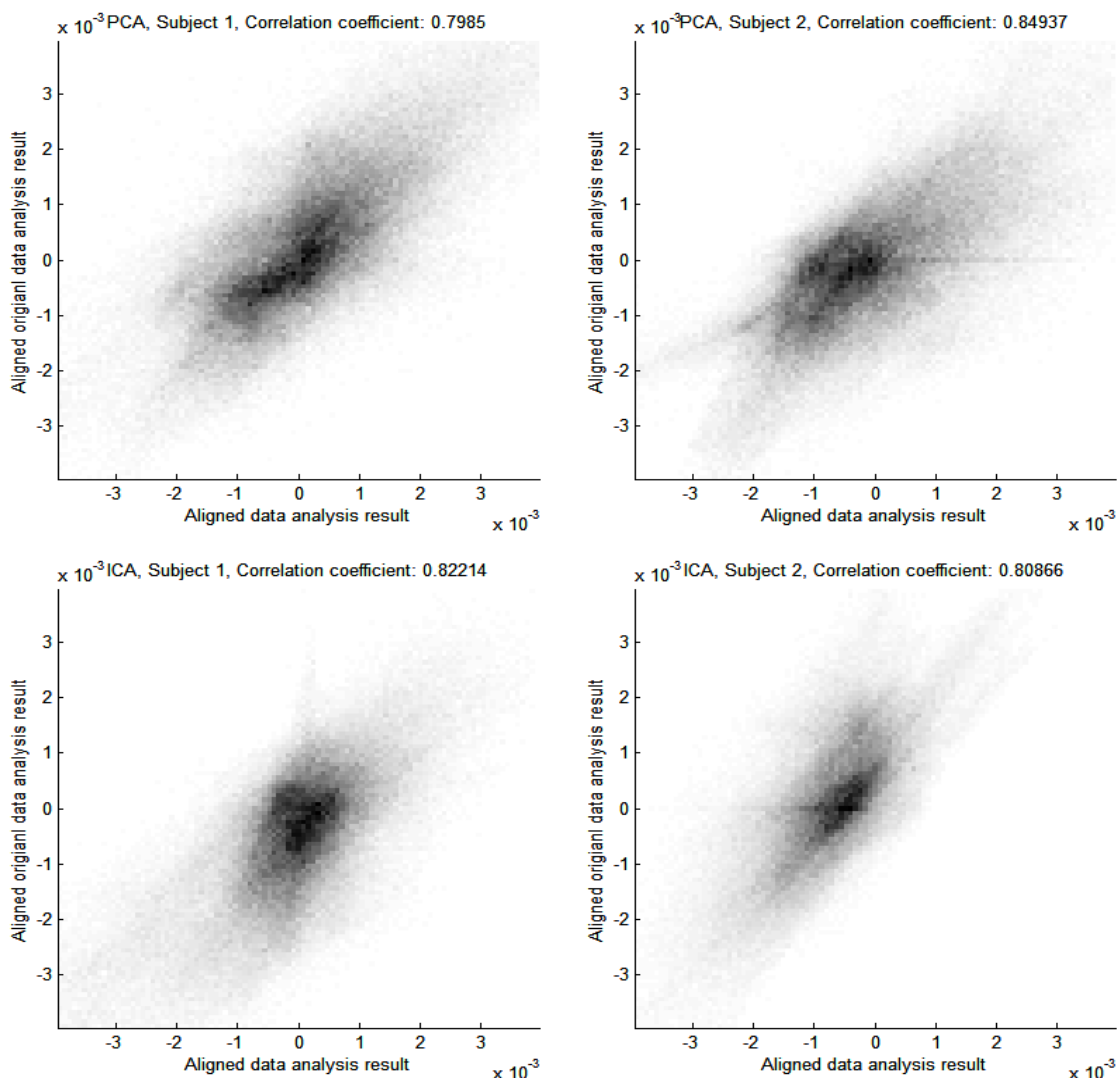
Fig. 4. Two-dimensional histograms comparing the spatial patterns obtained from the original and spatially aligned data analyses. The horizontal axes corresponds to the voxel values obtained from the analysis of spatially aligned data. The vertical axis correspond to the voxel values of the volumes obtained by spatially normalizing the results of original data analysis.

subject was deliberately misaligned using an affine transformation. Temporal patterns corresponding to the 4th principal component have the highest correlation with the on-off experimental paradigm.

To compare the spatial patterns obtained with and without spatial normalization we again used the SPM software to normalize the original data results to the same spatial template. The voxel values obtained in this way together with the voxel values obtained by the analysis of the aligned data are used to construct the two-dimensional histograms shown in Fig. 4 These histograms again show a high degree of correlation: 0.7985 (subject 1) and 0.85851 (subject 2) for the PCA and 0.82214 (subject 1) and 0.80866 (subject 2) for the ICA. In Fig. 5 and Fig. 6 we show the "pre-alignment" spatial patterns obtained from the spatially normalized data and "post-aligned" spatial patterns obtained by spatially normalizing the results of the original data analysis. These images clearly show an activation in the occipital lobe.

VI. CONCLUSION

In this paper we investigated the effects of spatial alignment transformations on PCA and ICA image analysis. We found simple conditions under which the patterns obtained by PCA and ICA are essentially unaffected by spatial transformations, meaning that spatial normalization of images may not be necessary, or can be applied after analysis if a common coordinate system is desired for the spatial patterns.

Our analysis can also be interpreted in a different way. It can be argued that analysis without pre-alignment of the images may be just as valid as the standard approach, because it is based on the original data, and does not depend on the accuracy or dependability of any alignment algorithm. Nevertheless, our preliminary experimental results appear to indicate that the results obtained with and without alignment are virtually identical.

In future work, we will consider the effects of more-complicated warping procedures across a wider range of data
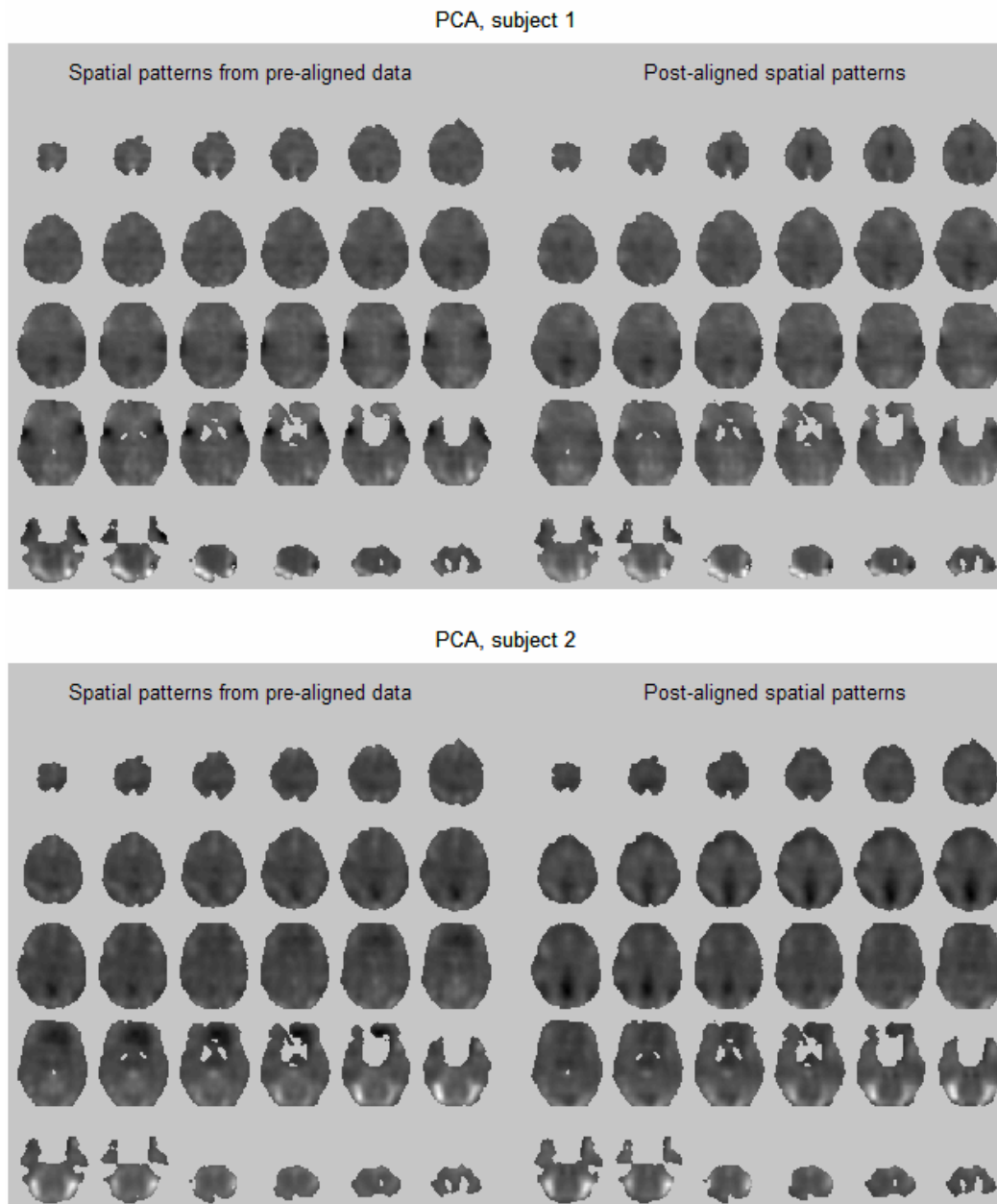
Fig. 5. The spatial patterns obtained from the PCA analysis of the original and spatially normalized (aligned) data for both subjects. The spatial patterns on the left are obtained from the analysis of spatially aligned data. The spatial patterns on the right are obtained by spatially normalizing the results of the original data analysis.

sets to further verify our conclusions. We will also study application of this approach to MEG, where spatial alignment of the original data can be difficult.

APPENDIX

DERIVATION OF THE ONE-STEP TEMPORAL MOLGEDEY-SCHUSTER ICA ALGORITHM

In concept, the Molgedey-Schuster ICA algorithm consists of two steps: a whitening operation (PCA), followed by a rotation operation to induce independence of the components. Here we show that these two steps can be represented more simply in a single combined operation.

As before, let $\mathbf{Z} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$ denote the reduced SVD of $\mathbf{Z}$. Now let $\mathbf{Q}^T$ be the matrix operator that rotates the PCA basis vectors in $\mathbf{V}$ to transform them into ICA source vectors. Thus, the source matrix becomes

$$\mathbf{S} = \mathbf{Q}\mathbf{V}^T. \tag{A.1}$$

Equating the SVD of $\mathbf{Z}$ with the ICA of $\mathbf{Z}$, i.e., $\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T = \mathbf{A}\mathbf{S}$, it immediately follows that the mixing matrix is given by
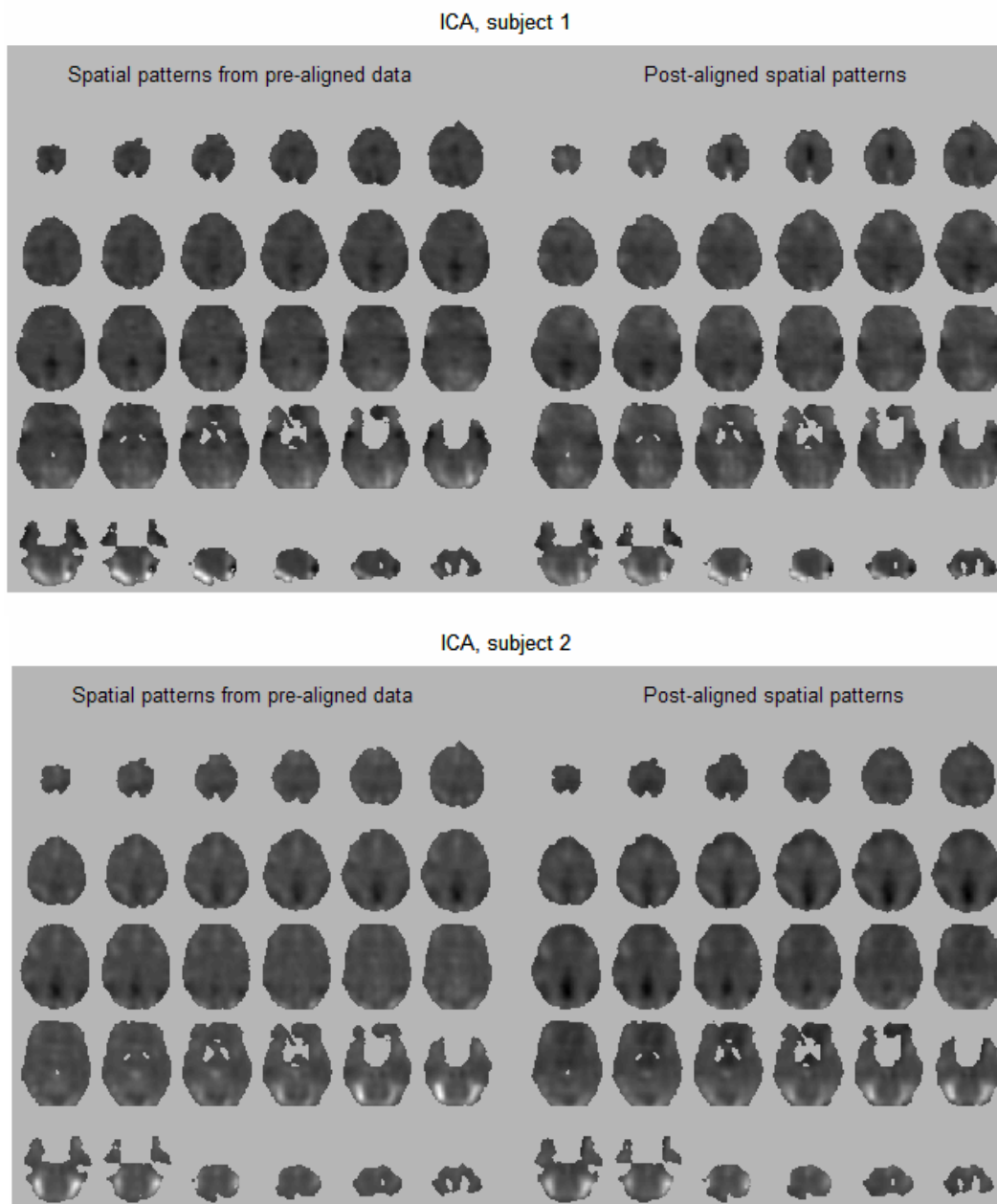
Fig. 6. The spatial patterns obtained from the ICA analysis of the original and spatially normalized (aligned) data for both subjects. The spatial patterns on the left are obtained from the analysis of spatially aligned data. The spatial patterns on the right are obtained by spatially normalizing the results of the original data analysis.

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{Q}^{T} . \tag{A.2}$$

From (A.1) it immediately follows that the time-lagged autocorrelation matrix of the sources is given by:

$$\mathbf{C}_{s}(\tau) = \mathbf{Q}\mathbf{C}_{v}(\tau)\mathbf{Q}^{T} . \tag{A.3}$$

In the Molgedey-Schuster algorithm, the aim is to enforce independence by causing the time-lagged cross correlation of the sources to be zero, i.e., $\mathbf{C}_{s}(\tau)$ must be diagonal. Here the data are assumed to be statistically stationary. Therefore,

(A.3) is an eigenvector equation, which defines the orthogonal matrix $\mathbf{Q}^{T}$ as the eigenvector matrix of $\mathbf{C}_{v}(\tau)$, i.e.,

$$\mathbf{C}_{v}(\tau)\mathbf{Q}^{T} = \mathbf{Q}^{T}\mathbf{C}_{s}(\tau) . \tag{A.4}$$

From $\mathbf{Z} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^{T}$, we have $\mathbf{V} = \boldsymbol{\Lambda}^{-1}\mathbf{U}^{T}\mathbf{Z}$. Thus, the time-lagged autocorrelation matrices of $\mathbf{V}$ and $\mathbf{Z}$ are related by:

$$\mathbf{C}_{v}(\tau) = \boldsymbol{\Lambda}^{-1}\mathbf{U}^{T}\mathbf{C}_{z}(\tau)\mathbf{U}\boldsymbol{\Lambda}^{-1} . \tag{A.5}$$

Substituting (A.5) into (A.4), we obtain:

$$\mathbf{\Lambda}^{-1}\mathbf{U}^{T}\mathbf{C}_{Z}(\tau)\mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{Q}^{T} = \mathbf{Q}^{T}\mathbf{C}_{S}(\tau)$$

$$\mathbf{C}_{Z}(\tau)\mathbf{U}\mathbf{\Lambda}^{-2}\mathbf{U}^{T}(\mathbf{U}\mathbf{\Lambda}\mathbf{Q}^{T}) = (\mathbf{U}\mathbf{\Lambda}\mathbf{Q}^{T})\mathbf{C}_{S}(\tau) \qquad (A.6)$$

$$\mathbf{C}_{Z}(\tau)\left(\mathbf{U}\mathbf{\Lambda}^{-2}\mathbf{U}^{T}\right)\mathbf{A} = \mathbf{A}\mathbf{C}_{S}(\tau)$$

The last equation in (A.6) is an eigenvector equation that can be used to calculate the mixing matrix **A** directly in one step.

REFERENCES

[1] S.C. Strother, "Evaluating fMRI preprocessing pipelines," *IEEE Eng. Med. Biol. Mag.* (in press).

[2] M. Svensén, F. Kruggel, and H. Benali, "ICA of fMRI group study data," *Neuroimage*, vol. 16, pp. 551-563, 2002.

[3] A.S. Lukic, M.N. Wernick, L.K. Hansen, J. Anderson, and S.C. Strother, "A spatially robust ICA algorithm for multiple fMRI data sets," *Proc. IEEE Intl. Symp. Biomed. Imaging*, pp. 839-842, 2002.

[4] V.J. Schmithorst and S.K. Holland, "Comparison of three methods for generating group statistical inferences from independent component analysis of functional magnetic resonance imaging data," *J. Magn. Reson. Imaging*, vol. 19, pp. 365-368, 2004.

[5] C.F. Beckmann and S.M. Smith, "Tensorial extensions of independent component analysis for multisubject fMRI analysis," *NeuroImage* vol. 25, pp. 294-311, 2005.

[6] V. D. Calhoun, T. Adali, V. B. McGinty, J. J. Pekar, T. D. Watson, G. D. Pearlson, "fMRI activation in a visual-perception task: Network of areas detected using the general linear model and independent components analysis", *Neuroimage*, vol. 14, pp. 1080-1088, 2001

[7] V. D. Calhoun, T. Adali, G. D. Pearlson, J. J. Pekar, "A method for making group inferences from functional MRI data using independent component analysis", *Human Brain Mapping*, vol. 14, pp. 140-151, 2001

[8] K. Friston, J. B. Poline, A. P. Holmes, C. D. Frith, and R. S. J. Frackowiak, "A multivariate analysis of PET activation studies," *Human Brain Mapping*, vol. 4, pp. 140–151, 1996.

[9] U. Kjems, S.C. Strother, J.A. Anderson, I. Law, L.K. Hansen, "Enhancing the multivariate signal of [^{15}O]water PET studies with a new non-linear neuroanatomical registration algorithm," *IEEE Trans. Med. Imaging.* vol. 18, pp. 306-319, 1999.

[10] S.C. Strother, S. LaConte, L.K. Hansen, J. Anderson, J. Zhang, S. Pulapura, and D. Rottenberg, Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics. I. A preliminary group analysis," *Neuroimage* vol. 23S1, pp. S196-S207, 2004.

[11] J.E. Jackson, *User's Guide to Principal Components.* New York: Wiley, 1991.

[12] A. Kyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis.* New York: Wiley, 2001.

[13] L. Molgedey and H. Schuster, "Separation of independent sources using time-delayed correlations," *Phys. Rev. Lett.*, vol. 72, pp. 3634-3637, 1994.

[14] J. Larsen, L.K. Hansen, and T. Kolenda, "On Independent Component Analysis for Multimedia Signals," in *Multimedia Image and Video Processing*, L. Guan, S. Y. Kung, and J. Larsen, Eds. New York: CRC Press, pp. 175-200, 2002.

[15] K.S. Petersen, L.K. Hansen, T. Kolenda, E. Rostrup, and S.C. Strother, "On the independent components of functional neuroimages," *Proc. ICA-2000*, pp. 615-620, 2000.

[16] R. P. Woods, S. T. Grafton, J. D. G. Watson, N. L. Sicotte, and J. C. Mazziotta, "I: Intersubject validation of linear and nonlinear models," J. Comput. Assist. Tomogr., vol. 22, no. 1, pp. 153–165, Jan. 1998.

[17] T. Talairach and J.A. Tornoux, *Stereotactic Coplanar Atlas of the Human Brain*. Stuttgart: Thieme, 1988.

[18] S. LaConte, J. Anderson, S. Muley, J. Ashe, S. Fruitiger, K. Rehm, L.K. Hansen, E. Yacoub, X. Hu, D. Rottenberg, and S. Strother, "The evaluation of preprocessing choices in single-subject BOLD-fMRI using NPAIRS performance metrics," *Neuroimage*, vol. 18, pp. 10-27, 2003.