COMPARISON OF TEMPORAL AND SPATIAL ICA IN FMRI DATA ANALYSIS

Silke Dodel, J. Michael Herrmann, Theo Geisel

Max Planck Institut für Strömungsforschung
Bunsenstraße 10, 37073 Göttingen, Germany

ABSTRACT

Independent Component Analysis (ICA) has recently drawn attention also in analyzing data from functional Magnetic Resonance Imaging (fMRI). fMRI is a promising method to determine noninvasively the spatial distribution of brain activity in a given situation, e.g. in response to a given stimulus. ICA reflects the underlying statistically independent processes of the activity.

We compare the results of applying both, temporal and spatial ICA, on the data of a motor task experiment. It turns out that none of the two ways is a priori superior over the other, so in order to get fuller insight into the activity distribution both should be applied and their implications be determined from the features of interest in the results.

1. INTRODUCTION

Functional Magnetic Resonance Imaging (fMRI) in the recent years has become a widespread tool to determine noninvasively the spatial distribution of brain activity in response to a given stimulus. It is an indirect method in that the signal is dependent on the blood oxygen level, which is changing in response to neural activity rather than on the neural activity itself. Though this makes the temporal resolution rather low, the spatial resolution of the brain images is in the order of 1 mm, which is much better than for other noninvasive measurement methods of the brain activity as e.g. Electroencephalography (EEG) or Magnetoencephalography (MEG). However, often the functional cannot be identified from the plain image data as the relative signal intensity is rather small, so data analysis methods are applied to determine the activity distribution. Independent component analysis (ICA) has hereby recently drawn attention as a more powerful analysis method than the widely used correlation analysis in which a predefined stimulus time course is correlated with each pixel time course. Although this correlation analysis provides a simple and easy interpretable way to determine the stimulus related activity, it suffers from the fact, that prior assumptions have to be made as about the shape of the stimulus time course, before applying the method. Furthermore, activity which is not related to the stimulus time course cannot be detected. ICA provides a method to extract multiple components of brain activity based on the assumption that the underlying brain processes are statistically independent and the data is a linear superposition of these processes. Thereby the assumption of statistical independence can be made in either temporal and spatial respect. Temporal independence refers to the time courses of the processes underlying the data whereas spatial statistical independence refers to the activity distributions in the brain images corresponding to the underlying processes.

The paper is organized as follows: First, the method of ICA is introduced, then the motor task experiment is described and finally the results of applying temporal and spatial ICA to the data are shown and discussed.

2. ICA

Independent component analysis is based on the assumption, that the data is a linear superposition of statistically independent processes. With this assumption, the data can be written as

$$X = AS, \tag{1}$$

where X is the $m \times k$ data matrix, A is a constant $m \times m$ mixture matrix and S the $m \times k$ matrix of the independent components. Both, A and S, have to be determined in ICA. In fMRI each column of X corresponds to an image consisting of m pixels and each row of X corresponds to the time course of a pixel. One can thus either consider the data as an ensemble of images, i.e. as a set of vectors in an m-dimensional data space or as an ensemble of time courses, corresponding to vectors in a k-dimensional data space. Obviously the data span a subspace of dimension equal to rank X in both cases. In temporal ICA the rows of S are assumed to constitute statistically independent time courses a linear mixture of which is measured. The linear mixture is given by the matrix A. Similarly in spatial ICA the data is assumed to be a linear mixture of the columns of S, which are assumed to constitute statistically independent images.

Statistical independence means that the joint probability density p of respectively each column or row $\mathbf{s}_j$ of the matrix S can be written as

$$p(\mathbf{s}_j) = \prod_{i=1}^{r} p_i(s_{ij}), \tag{2}$$

where p_i are the marginal probability densities of the elements of the row or column $\mathbf{s}_j$. $r = m$ or $r = k$ depending on temporal or spatial ICA being considered. It is not always possible to find ICs $\mathbf{s}_j$ such that eq. (2) holds, thus often ICs are determined by minimizing a 'distance' measure $D(p(\mathbf{s}_j), \prod_{i=1}^{r} p_j(s_{ij}))$ both sides of the equation. Prior to the actual ICA most algorithms apply Principal Component Analysis (PCA), an orthogonal coordinate transform of the data with respect to which the data is uncorrelated and which reflect the directions of the most prominent variances in the data. This is achieved by diagonalizing the covariance matrix of the data. The eigenvectors then form the new coordinate system. With PCA the data is made statistically independent up to the second order before with ICA the independence is increased to higher statistical orders. In case of jointly gaussian distributed data, where the moments of higher than second order equal zero, PCA is equivalent to ICA. Unlike for the Principal Components (PCs) of the data, which are ordered according to the variance they reflect, no intrinsic order exists for the Independent Components (ICs), since a row permutation of S together with a column permutation of A does not change X. Also the amplitudes of the rows of S can be varied without affecting X by scaling A accordingly.

We used the algorithm described in [2], which diagonalizes cumulant matrices obtained from an estimation of the joint probability density p. Using kurtosis, i.e. approximating p only to the fourth order, it remains uncertain, whether the data are separated correctly, even if it is possible in principle. However, usually there are not enough data to consistently estimate the underlying probability density p to a higher order, so this approach seems to be reasonable.

ICA algorithms based on other statistical concepts can be found in [1], [3] and [4]. An overview of different methods to tackle ICA can be found in [5]. Spatial ICA was applied to fMRI measurements in [6], but was not compared to temporal ICA.

3. THE EXPERIMENT

The experiment consisted of two parts each taking 280 s (cf. Fig. 1). In the first part after a 40 s rest the subject was asked to tap the fingers of both hands for 20 s and then again rest for 20 s. This finger tapping cycle of 20 s tapping and 20 s rest was repeated six times. The second part of the experiment was identical to the first except that the subject was not supposed to actually do the finger tapping but only to imagine it according to the same time course as in the second part.

The images were taken with a frequency of 2 Hz, allowing a frequency resolution up to 1 Hz (the finger tapping frequency of $2 - 3$ Hz could hence not been resolved, only the finger tapping *cycle* frequency of $\frac{1}{40s} =$ 0.025 Hz). Each image vector contained $m = 4015$ pixels arranged as a 55×73 matrix and $k = 560$ images were taken. In the analysis, the first 20 images, corresponding to the first 10 s were omitted for magnetization steady state reasons.

4. RESULTS

Temporal and spatial ICA was performed for the data of each of the two parts seperately. Before applying ICA the dimension of the data was reduced by projecting it onto its first 8 principal components. Principal Component Analysis can also be performed either temporally or spatially by diagonalizing the temporal or spatial covariance matrix, respectively. We applied both, temporal and spatial PCA for dimension reduction. 8 principal components is a reasonable choice since the higher principal components were rather unstructured and could therefore be assigned to noise.

To each of the reduced data matrix both temporal and spatial ICA was applied, resulting in 4 sets of 8 Independent Compents, respectively.

In the temporal case the independent components are time courses and in the spatial case activity distributions. However, for simplicity, we will refer in both cases to the activity distribution as to the independent component, since each activity distribution can be assigned a time course and vice versa. The time course is reflecting the contribution of the activity distribution to the data at each time instant.

We computed the power spectrum of each time course, normalized to the total power to allow comparison between different spectra. The independent components with the highest frequency contribution at the finger tapping cycle frequency of 0.025 Hz was chosen from each set of independent components, because it reflects best the activity in response to the external finger tapping and is therefore most suitable for comparison. The result for the actual finger tapping (part one of the experiment) is shown in Fig. 2 and Fig. 3 shows the result for the imaginary finger tapping. In each row is shown the activity distribution (the front side of the brain is directed to the left, the view is from below), the time course of the activity distribution and the power spectrum of the time course. As result of the actual finger tapping an increased symmetrically bilaterally distributed activity in the motor cortex reflecting the stimulus time course is expected. Activity in the motor cortex can be seen in Fig. 2 (b) (blue area on the lower right hand side). However, the corresponding contralateral activity on the other half of the brain is only weak indicating that the image plane did not cut both regions equally during the measurement. For the imaginary finger tapping activity in the supplementary motor cortex is expected which is seen in Fig. 3 (a) the blue area on the middle to the left hand side of the activity distribution.

It turned out, that the spatial and temporal PCA were essentially identical except for the first component, which reflected the mean activity distribution of the data. The first spatial PC shows a more "brain-like" structure than the rest of the components, which influences also the result of the ICA as can be seen in Fig. 2 (c), where temporal ICA was performed after spatial PCA. The corresponding result for spatial ICA does not show a similar prominent brain-like shape, because while performing ICA the mean is substracted, which is in this case the spatial mean, thus the shape cancelles out. The combination of spatial PCA and spatial ICA shows a rather irregular time course without reflecting more spatial structure and thus seems not to be suited as tool for the analysis. The temporal coherence of the underlying processes of the data should thus be taken into account.
To determine how well each of the PCA-ICA combinations could separate the stimulus related activity they were compared to a control set of ICs computed by temporal and spatial ICA from the projection of the actual and imaginary finger tapping data onto the 4 or 3 temporal PCs, respectively, showing the most prominent contribution of the finger tapping cycle frequency (projecting onto the corresponding spatial PCs to build a control set would have made no difference because of the similarity of spatial and temporal PCs with exception of the first spatial PC which did not contribute here). It turned out that the combination of temporal PCA with temporal and with spatial ICA, respectively, had a visually identical corresponding component in the respective control set for the actual finger tapping and similarly for the imaginary finger tapping the combination temporal PCA - spatial ICA. The combination temporal PCA - temporal ICA differed slightly from the corresponding component of the control set but was still very similar.

5. CONCLUSION

From the findings of this experiment a projection of the data to spatial principal components prior to the application of ICA is not recommended. For the combination with the projection onto temporal PCs the results suggest slight superiority of spatial ICA. However, the difference is not prominent enough to impose a decision. Thus it seems that neither spatial nor temporal ICA is a priori superior over the other. In order to get fuller insight into the activity distribution both should be applied after temporal PCA and their implications be determined from the features of interest in the results. We thank Peter Fransson and Peter Dechent from the MPI for Biochemistry in Göttingen for very stimulating discussions and for providing us with the fMRI data of the finger tapping experiments.

6. REFERENCES

[1] J. Bell, T. J. Sejnowski, An Information-Maximization Approach to Blind Separation and Blind Deconvolution, 1995, *Neural Computation*, **7**, 1129-1159

[2] J. F. Cardoso, A. Souloumiac, 1993, Blind Beamforming for Non Gaussian Signals, *IEE-Proceedings-F*, **140**, N^o 6, 362-370

[3] P. Comon, 1994, Independent component analysis, A new concept?, *Signal Processing*, **36**, 287-314

[4] A. Hyvärinen, 1997, Independent component analysis by Minimization of Mutual Information, *Report A46*, Otaniemi

[5] T. W. Lee et al., 1999, A Unifying Information-theoretic Framework for Independent Component Analysis, *International Journal of computers and mathematics with applications*, (in press)

[6] M. J. McKeown et al., 1998, Analysis of fMRI Data by Blind Separation Into Independent Spatial Components *Human Brain Mapping*, **6**, 160-188

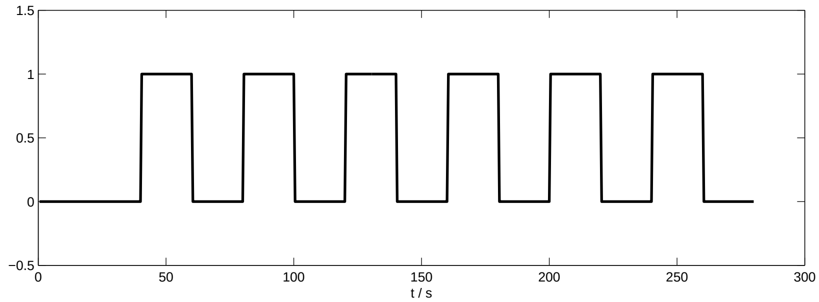

Figure 1:
Stimulus time course in the first part of the experiment. The value of 1 indicates 'stimulus on', 0 indicates 'stimulus off'.

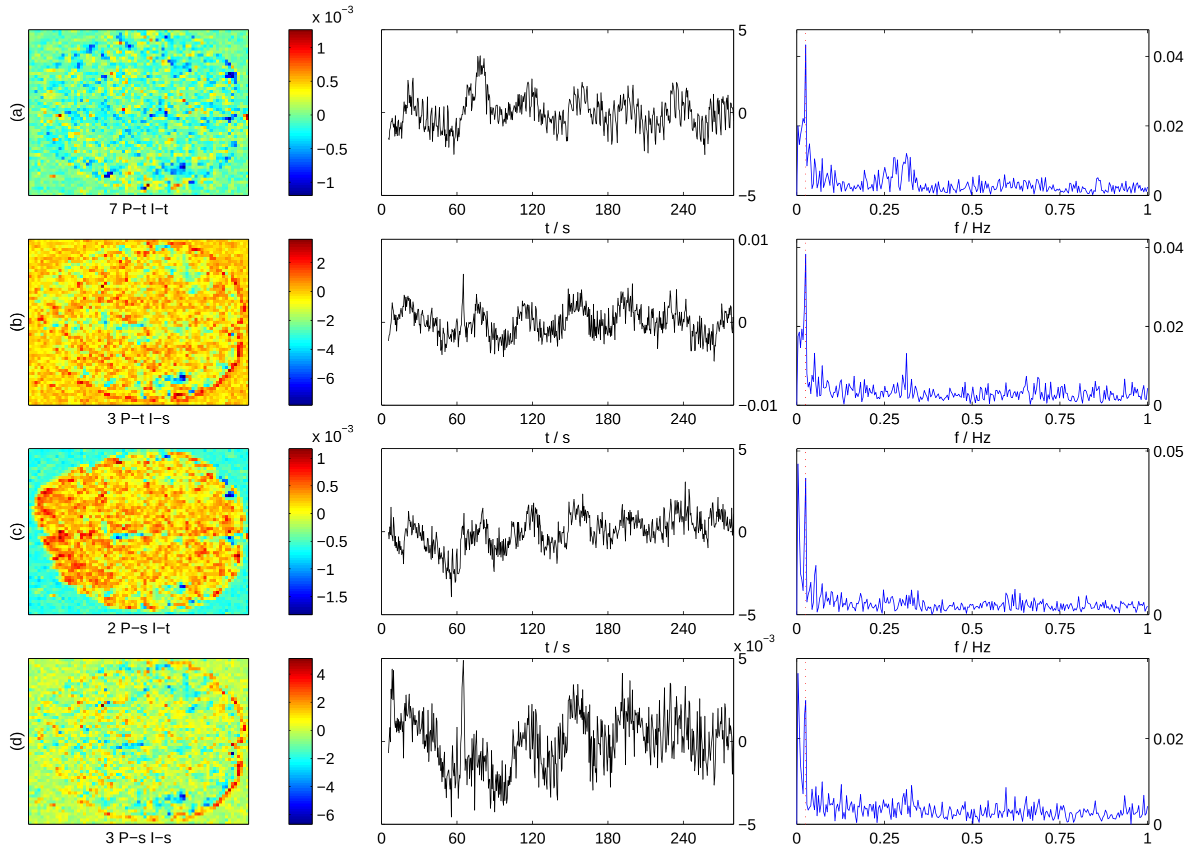

Figure 2:
Actual finger tapping:
Results with the highest frequency contribution present at 0.025 *Hz (red dotted curve in the power spectrum depicted in the right column). Middle column: time courses (arbitrary units on the ordinates). Left column: Activity distributions.*
P-t: temporal PCA, P-s: spatial PCA, I-t: temporal ICA, I-s: spatial ICA

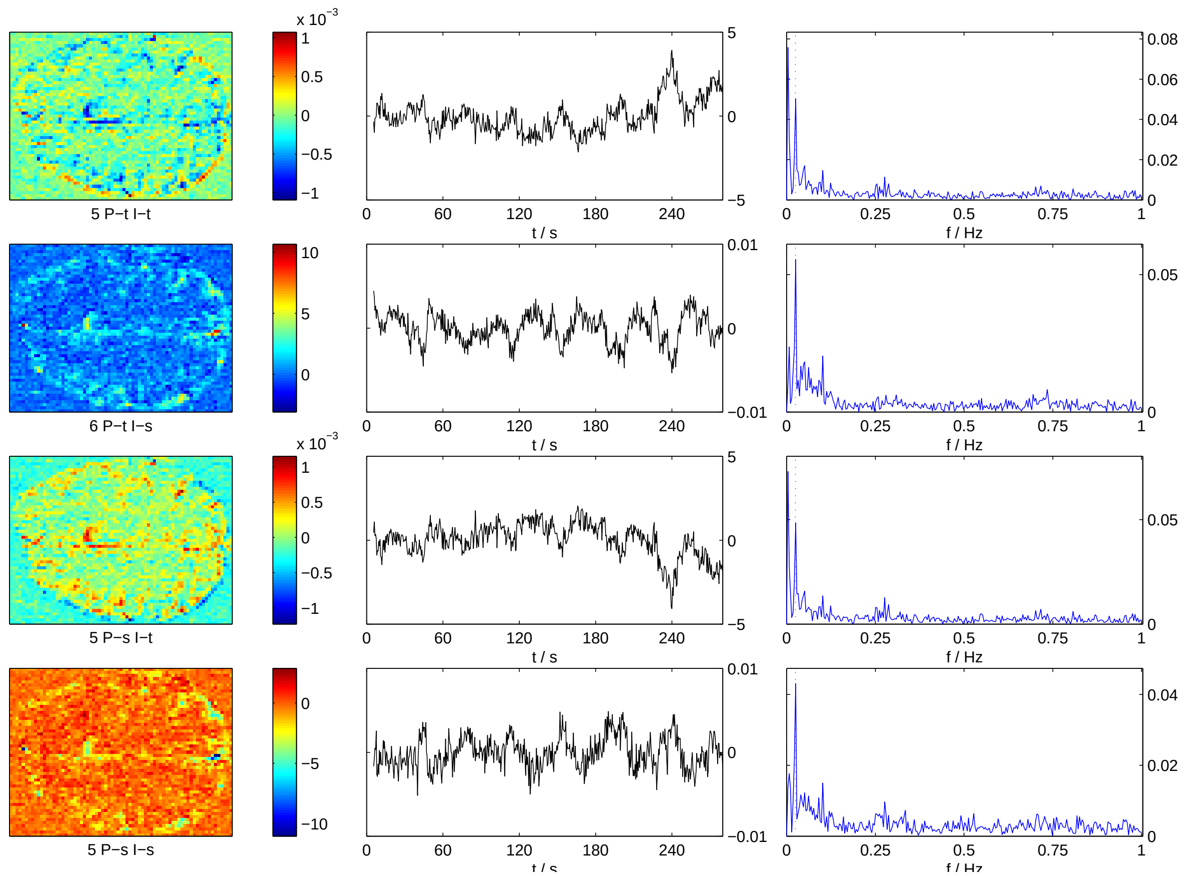

Figure 3:
Imaginary finger tapping:
Results with the highest frequency contribution present at 0.025 *Hz (red dotted curve in the power spectrum depicted in the right column). Middle column: time courses (arbitrary units on the ordinates). Left column: Activity distributions.*
P-t: temporal PCA, P-s: spatial PCA, I-t: temporal ICA, I-s: spatial ICA