ON THE INDEPENDENT COMPONENTS OF FUNCTIONAL NEUROIMAGES

K.S. Petersen, L.K. Hansen, T. Kolenda

THOR Center for Neuroinformatics
Dept. Mathematical Modelling,
Technical University of Denmark, B321
DK-2800 Lyngby, Denmark

E. Rostrup[a], S. C. Strother[b]

a. DRCMR, Hvidovre Hospital,
Hvidovre, Denmark.
b. PET Imaging Service, MVAMC
Minneapolis, Minnesota, USA

ABSTRACT

We discuss the application of ICA procedures to fMRI (functional Magnetic Resonance Imaging) sequences. While principal component analysis can identify activation patterns that are uncorrelated in both space and time ICA can identify events that are independent in either time or space. We show that the activation related components found by either spatial or temporal independency are consistent, hence robust to choice of spatial or temporal separation and to choice of ICA approach. We discuss these issues in the context of three ICA algorithms applied to an fMRI visual activation study.

1. INTRODUCTION

Functional magnetic resonance imaging (fMRI) is based on the so-called BOLD (blood oxygen level dependent) contrast, namely that blood perfusion in the brain has a measurable effect on the magnetization relaxation time. In an fMRI activation study, healthy volunteers are maintained in controlled mental states (typically a baseline control state and an active state, the "task"). The fMRI signal is an indirect measurement of neural activation, though the *hemodynamic response*: the characteristic blood perfusion changes following neural activity. The typical local signal change is below 5%. For additional detail about brain activation studies and the fMRI technique, see [20].

The fMRI signal has many sources. On top of "neuronal noise", i.e., uncontrollable activations, the fMRI signal is confounded by several physiological signal components the most prominent being the heart beat at about 1 Hz, and the breathing signal at about 0.1-0.2 Hz. Most fMRI data sets are acquired at sampling frequencies 0.2-0.5 Hz, hence the heart signal is aliased and not represented faithfully, in this work we analyse a dataset sampled at 3 Hz thus avoiding these problems.

Whether the heart signal is aliased or not, the fMRI signal is of a complicated spatio-temporal nature and is consequently approached by many different signal processing strategies see e.g. [19, 13]. Some strategies emphasize spatial other temporal aspects. Analysis of single pixel time series say by cross-correlating with a stimulus indicator reference function (equivalent to a t-test) [4] is probably the most widespread analysis method. However, the high degree of connectivity in the brain suggests that one should expect global correlation in response to most types of stimulation. Global spatial correlations can be probed by principal component analysis (PCA). PCA was introduced in functional neuroimaging in [17]. More general global dependencies can be probed by independent component analysis (ICA) based on higher order or spatio-temporal statistical properties, see [14] for a review. With ICA we could hope to separate the different sources of the fMRI signal.

The basic tool for PCA is singular value decomposition (SVD). Let $X = [X_{m,n}]$ be a matrix of N columns each holding an fMRI image of M pixels. Most fMRI experiments produce short sequences ($N < M$). Using SVD X is decomposed as

$$X = U\Lambda V^\top, \quad X_{m,n} = \sum_\lambda U_{m,\lambda}\Lambda_\lambda V_{n,\lambda} \qquad (1)$$

where U is a $M \times N$ matrix, while Λ, V are $N \times N$ matrices. Λ is a diagonal matrix with non-negative ordered elements (the largest first)– the singular values. U are eigenvectors corresponding to the N non-zero eigenvalues of $XX^\top$, while V is an orthogonal matrix holding the eigenvectors of $X^\top X$. Viewed as temporal sequences, $\Lambda V^\top$ are "uncorrelated source signals". The associated spatial patterns U are orthogonal being eigenvectors of a symmetric real matrix.

However, as noted above we are often interested in a slightly more general separation of image sources that are independent in time, but not necessarily orthogonal in space, *or* independent in space but not necessarily uncorrelated in time.

McKeown et al. analysed an fMRI data set arguing for spatial independence and found good separation of modes that were task related transiently task related as well as confounding modes representing e.g., head motion [16]. Biswal and Ulmer analysed fMRI data searching for temporal independent activation sequences, however limiting the spatial complexity to relatively small regions (M=30 pixels) [6]. They found that the temporal ICA analysis was able to resolve two different induced effects on the fMRI signal, a task induced effect and CO_2 inhalation (hypercapnia). Since hypercapnia induce a global effect of enhanced BOLD signal, searching for spatial independence is not so obvious. Lange et al. [13] compared spatial ICA (based on the Bell-Sejnowski algorithm) on simulated and real fMRI. They found that ICA could identify features in brain maps not accessible by simple correlation or t-test based methods.

In this communication we will compare the two different ways of analysing fMRI using ICA for a single task visual study.

The ICA decomposition can be written

$$X = AS, \quad X_{m,n} = \sum_{\mu} A_{m,\mu} S_{\mu,n} \tag{2}$$

with A being a matrix of image columns and S a matrix of time series rows. Spatial ICA then means searching for independent image vectors while temporal ICA means searching for independent rows in S

For short image sequences we can make use of the SVD for an initial simplification of the ICA problem. The approach taken here is similar to the so-called "Cure for Extremely Ill-posed Learning" [12] used to simplify supervised learning in short image sequences, and is based on a projection of the image vectors onto a basis spanned by the observed images.

If we want to perform temporal ICA the following data reduction step is possible. Let $X = U\Lambda V^\top$, then

$$Y \equiv U^\top X = U^\top AS \equiv BS, \tag{3}$$

with Y being an $N \times N$ matrix. Finding B, S using an ICA algorithm A can be reconstructed using U. The columns of A will then be "eigenimages" dual to the independent time series in S. If we on the other hand want to perform spatial ICA then we can use the tranformation,

$$Y^\top \equiv V^\top X^\top = V^\top S^\top A^\top \equiv (BS)^\top. \tag{4}$$

Again, by projection using the eigen time series V we can recover the full time series S from a reduced representation.

This transformation is a simple rewriting of the separation problem and does not introduce any loss of generality. However we note that it often may be possible to further limit the dimensionality of the SVD subspace retained in Eq. (4) at the price of an information loss.

The ICA decomposition in the reduced space can be achieved by a variety of algorithms with different assumptions. We have chosen to compare the original Bell and Sejnowski algorithm [5], the decorrelation based method of Molgedey and Schuster [18], and Attias and Schreiner's dynamic component method [2, 3].

The three ICA methods can be applied in "spatial" and "temporal" modes.

The Bell-Sejnowski learning algorithm can be viewed as a maximum-likelihood method assuming linearly mixed "white" source signals with a common density $p(s) \approx 1/\cosh(s)$ [5, 15]. The Molgedey-Schuster approach is based on decorrelation using that the independent sources have different autocorrelation functions (spatially or temporally) [18, 8, 10]. In [10] we proposed a minor modification of the Molgedey-Schuster method that avoids a problem with complex valued eigenvalues. Attias and Schreiner proposed a very rich ICA framework based on higher order statistics and decorrelation [2, 3], essentially combining the Bell-Sejnowski and Molgedey-Schuster approaches. Furthermore, they allowed for completely general and learnable source distributions. Here we have implemented the so-called DCA-IFT version of their approach with a fixed, parameter free, bi-exponential source density

$$p(s) \approx \exp(-|s|).$$

We experienced some difficulty in getting the more flexible scheme with a gaussian mixture source distribution to converge for the present data.

2. THE VISUAL ACTIVATION STUDY

A single fMRI slice holding 128×128 pixels and cutting through primary visual cortex was acquired with a time interval between successive scans of TR = 333 msec. A window of $M = 6794 = 86 \times 79$ pixel covering all of the brain of the particular slice was extracted for this analysis. This sampling frequency is small enough to allow faithful representation of the heart signal. Visual stimulation in the form of a flashing annular checkerboard pattern was interleaved with periods of fixation. A run consisting of 30 scans of fixation, 31 scans of stimulation, and 60 scans of post-stimulus fixation was repeated 10 times.

3. RESULTS

An initial SVD was performed on the $(N = 1210) \times (M = 16384)$ data matrix. Sixteen principal components were kept in the SVD, hence, spatial ICA analyses were carried out on the reduced 16×16384 data matrix, while temporal ICA analyses were carried out on the reduced 16×1210 matrix. Running spatial and temporal ICA with the three different algorithms thus creates six individual results each having 16 independent components with an image (referred to informally as "eigenimage" c.f., the eigen images in PCA) and an associated time series.

Bell-Sejnowski and Attias-Schreiner's algorithms were both trained with the natural gradient learning rule [1]. The execution times of the six different ICAs are very different. The fastest is the Molgedy-Schuster algorithm, which is constructive and basically amounts to solving a 16×16 eigenvalue problem. It took 0.3 seconds for both spatial and temporal modes. The Bell-Sejnowski algorithm is iterative and took on the same machine 4.2 seconds to converge in temporal and 130 seconds in spatial mode. Attias and Schreiner's DCA-IFT algorithm is iterative and took 394 seconds in temporal mode and 2424 seconds in spatial mode to converge.

The power spectra of the resulting 6×16 time series are shown in Figure 1.

All results show a marked peak at low frequencies corresponding to the on/off activity induced by visual stimulation in several components. Two more peaks are seen at frequencies around 1 Hz, reflecting the heart beat. The heart signals are confined in one or a few components for the *spatial* Bell-Sejnowski ICA, the *temporal* Molgedey-Schuster ICA and in both spatial and temporal versions of Attias and Schreiner DCA-IFT algorithm. We conjecture that the reason for the problems for temporal Bell-Sejnowski ICA is that it mixes in with the bi-modal stimulus signal which has rather strong negative kurtosis. The problem for the spatial Molgedey-Schuster approach we believe arises from the spatially localised nature of the heart beat component (it is basically confined to vessels). Since the spatial Molgedey-Schuster signal separation is based on differences in component spatial autocorrelation function, it can be confounded with spatial white noise signal components. On closer inspection we find that the temporal Bell and Sejnowski algorithm is mixing stimulus and heart beat signal in a two-dimensional subspace. In Figure 2 we show the geometric picture with the correct angles between spatial components. If we now run the Molgedey-Schuster algorithm on these two mixed signals the heart beat and stimulus signals separate completely.

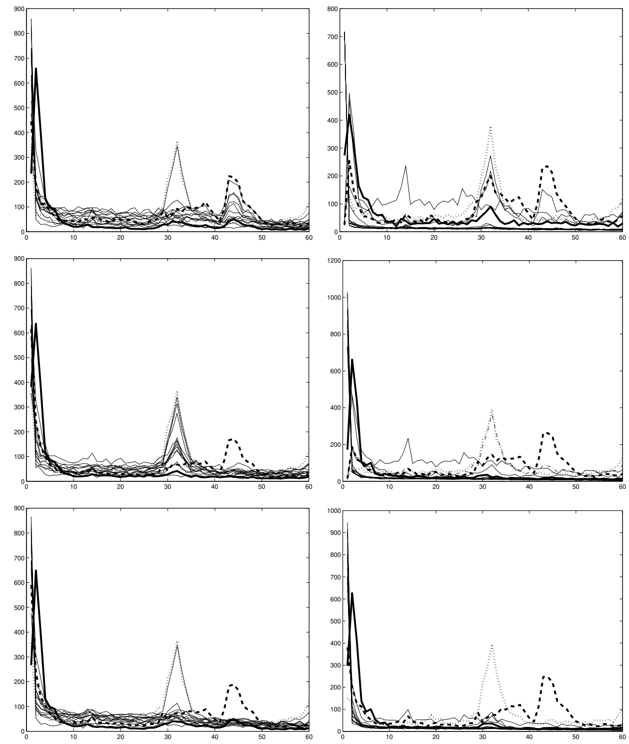

Figure 1: Power spectra obtained in independent component analyses of fMRI signals. $N = 1210$ fMRI images were acquired each holding $M = 6794$ pixels. The data matrix was reduced by PCA in either temporal or spatial dimensions to 16 principal components. Three different ICA algorithms (Bell-Sejnowski, Molgedey-Schuster and Attias-Schreiner) were run in spatial mode (data matrix 16×6794) or temporal mode (data matrix 16×1210). Spectra are shown for the 16 independent components identified by spatial (upper left) and temporal (upper right) Bell-Sejnowski ICA; spatial (mid left) and temporal (mid right) Molgedey-Schuster decorrelation, and spatial (lower left) and temporal (lower right) versions of Attias and Schreiner's DCA-IFT algorithms. All results show a marked peak at low frequencies corresponding to the on/off activity induced by visual stimulation in several components. Two more peaks are seen at frequencies around 1 Hz, reflecting the heart beat. The heart beat confounding peaks are confined in one or a few components for the *spatial* Bell-Sejnowski ICA, the *temporal* Molgedey-Schuster ICA and in both spatial and temporal versions of Attias and Schreiner DCA-IFT algorithm.

In Figures 3-5 we show the eigenimages corresponding to the three spatial (left panels) and three temporal (right panels) ICA's for the most task related component (measured by cross-correlation of the associated time sequence with an on/off stimulus reference function).

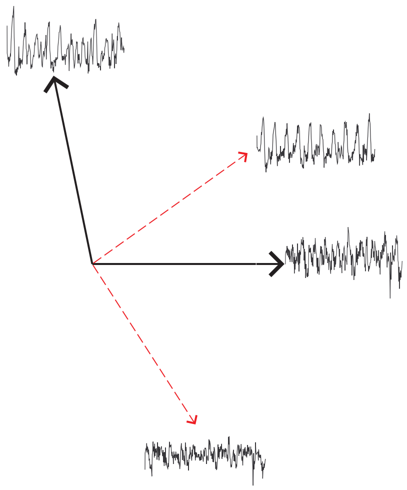

Figure 2: The temporal mode Bell-Sejnowski algorithm is mixing stimulus and heart beat signal in a two-dimensional subspace (full line arrows). We show the geometry in the two-dimensional subspace with the appropriate angles between spatial components. If we now run the Molgedey-Schuster algorithm on these two mixed signals the heart beat and stimulus signals separate completely (dashed arrows). The stimulus response signal shows ten characteristic peaks at activation.

We first notice the excellent overall agreement between the resulting "brain maps". This also holds true for the cases mentioned when the Bell-Sejnowski and Molgedey-Schuster algorithms have mixing problems in temporal and spatial separations respectively. Their dominant stimulus induced component are virtually identical to those found by Attias and Schreiner's algorithm. This allows us to conclude that there is a stimulus induced component which is very robust to ICA modeling assumptions, in this single task experimental design.

In Figure 6 we quantify the similarity between eigenimages found in spatial and temporal ICA. We have plotted the pixel values and the time series values respectively obtained in the two analyses with Attias and Schreiner's DCA-IFT algorithm. The correlation coefficients are $R_{\text{image}} = 0.86$ $R_{\text{time}} = 0.97$ respectively, strongly suggesting that we have found a robust separation of the task related component.

4. CONCLUSION

Brain imaging fMRI data was investigated using three different ICA approaches. While PCA results in orthogonal eigenimages and eigensequences, ICA can produce either independent spatial components *or* temporally independent components. Fortunately, we find that spatial and temporal ICA produce very similar

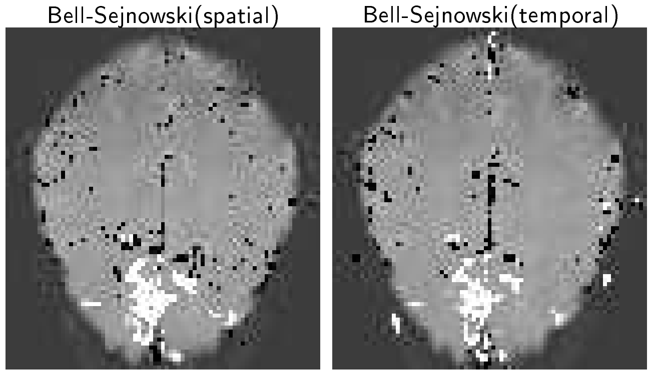

Figure 3: We show the eigenimages corresponding to the spatial (left panel) and temporal (right panel) Bell-Sejnowski ICA for the most task related component (measured by cross-correlation of the associated time sequence with an on/off stimulus reference function). These images are created by segmenting the eigenimage's 0.025 (black) and 0.975 (white)fractiles superimposed on the anatomical reference obtained by averaging all fMRI images. There is a marked activation in the lower mid area, corresponding to the primary visual cortex.

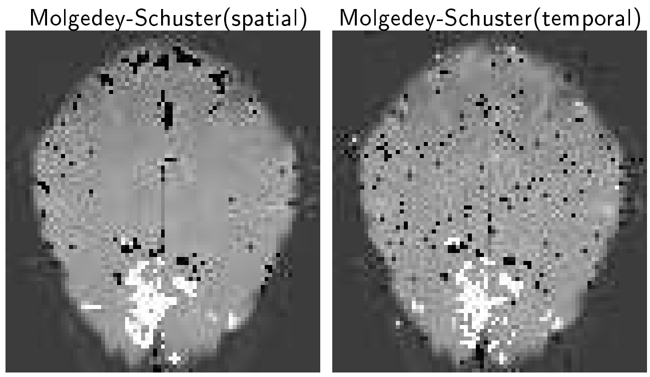

Figure 4: Task related eigenimage obtained by spatial (right) and temporal (left) ICA using the Molgedey-Schuster algorithm.

eigenimages ("brain maps"). For the Bell-Sejnowski approach we found some difficulty in separating the response signal and the confounding heart signal, when operated in temporal mode. This is most likely due to the bi-modal (negative kurtosis) nature of the response to on/off stimuli. The Molgedey-Schuster scheme, based on decorrelation, had difficulty separating in the spatial domain. This is most likely due to the similarity of the spatial autocorrelation for small vessel signals and spatial "white noise". The Attias-Schreiner approach essentially combines the Bell-Sejnowski and Molgedey-Schuster approach into one separation scheme that uses both decorrelation and non-linearity. Hence, this approach showed excellent separation in both spatial and temporal mode and indeed the brain maps and time se-

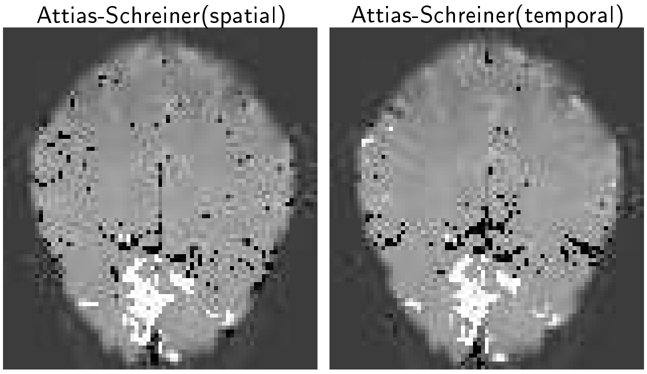

Figure 5: Task related eigenimage obtained by spatial (right) and temporal (left) ICA using Attias and Schreiner's algorithm.

ries obtained in the two modes were shown to correlate very well ($R_{\text{image}} = 0.86$ $R_{\text{time}} = 0.97$ respectively).

How general is the robustness of the activation pattern we see in this dataset? This is an interesting question for future research.

5. ACKNOWLEDGMENTS

This project is funded by the Human Brain Project P20 MH57180 "Spatial and Temporal Patterns in Functional Neuroimaging" and by the Danish Research Councils through the Danish Computational Neural Network Center, CONNECT and the THOR Center for Neuroinformatics.

6. REFERENCES

[1] S. Amari, A. Cichocki and H.H. Yang: *A New Learning Algorithm for Blind Signal Separation.* In Advances in Neural Information Processing Systems 8, Eds. D. Touretzky, M. Mozer, and M. Hasselmo, 757-763 (MIT Press Cambridge MA, 1996).

[2] H. Attias and C.E. Schreiner: *Blind Source Separation and Deconvolution by Dynamic Component Analysis.* In Neural Networks for Signal Processing VII: Proceedings of the 1997 IEEE Workshop, 456-465 (1997).

[3] H. Attias and C.E. Schreiner: *Dynamic Component Analysis* Neural Computation, **10** 1373-424 (1998)

[4] P.A. Bandettini, A. Jezmanowitcz, E.C. Wong, and J.S. Hyde: *Processing strategies for time course data sets in functiona MRI of the human brain* Magn. Reson. Med. **30**, 161-173, (1993).

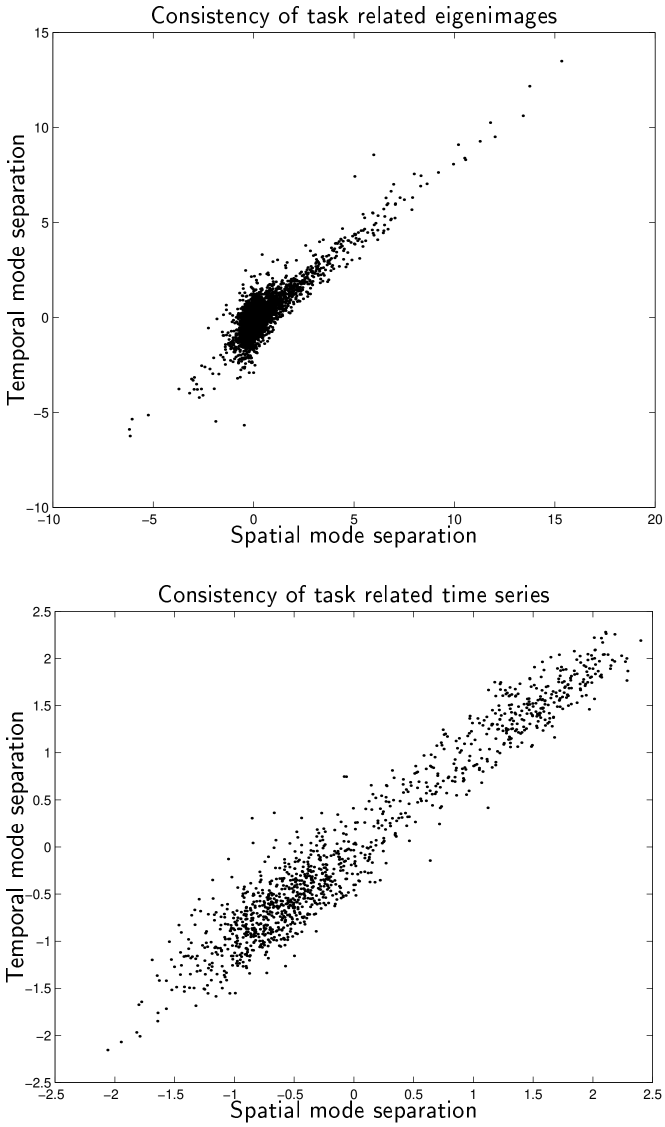

Figure 6: In Figures 3-5 we found very similar results from analysing in spatial and temporal modes. Here we show scatterplots of the images and time series found in the two modes by the Attias and Schreiner approach. The correlation coefficients are $R_{\text{image}} = 0.86$ $R_{\text{time}} = 0.97$ respectively, strongly suggesting that we have found a robust separation of the task related component.

[5] A. Bell & T.J. Sejnowski: *An Information-Maximization Approach to Blind Separation and Blind Deconvolution* Neural Computation **7**, 1129-59 (1995).

[6] B.B. Biswal amd J.L. Ulmer: *Blind Source Separation of Multiple Signal Sources of fMRI Data Using Independent Component Analysis* Journal of Computer Assisted Tomography **23** 265-271 (1999).

[7] P. Comon: *Independent Component Analysis: A New Concept?* Signal Processing, **36**, 287-314 (1994).

[8] L.K. Hansen and J. Larsen: *Source Separation in Short Image Sequences using Delayed Correlation* In Proceedings of the IEEE Nordic Signal Processing Symposium, Vigsø, Denmark 1998. Eds. P. Dalsgaard and S.H. Jensen, pp. 253-256, (1998)

[9] L.K. Hansen, J. Larsen, F.Å. Nielsen, S.C. Strother, E. Rostrup, R. Savoy, N. Lange, J.J. Sidtis, C. Svarer, O.B. Paulson: *Generalizable Patterns in Neuroimaging: How Many Principal Components?* NeuroImage **9**, 534-544 (1999).

[10] L.K. Hansen, J. Larsen and T. Kolenda *On Independent Component Analysis for Multimedia Signals* in L. Guan, S.Y. Kung and J. Larsen (eds.) Multimedia Image and Video Processing, CRC Press, to appear (2000).

[11] C. Jutten and J. Herault: *Blind Separation of Sources: An Adaptive Algorithm Based on Neuromimetic Architecture* Signal Processing **24**, 1-10 (1991).

[12] B. Lautrup, L.K. Hansen I. Law, N. Mørch, C. Svarer, S.C. Strother: *Massive weight sharing: A cure for extremely ill-posed problems.* In H.J. Hermanet al., eds. Supercomputing in Brain Research: From Tomography to Neural Networks. 137-148 World Scientific Pub. Corp. (1995).

[13] N. Lange, S.C. Strother, J.R. Anderson, F.Å. Nielsen, A.P. Holmes, T. Kolenda, R. Savoy, L.K. Hansen. *Plurality and Resemblance in fMRI Data Analysis* NeuroImage, **10** (3):282-303, (1999)

[14] T.-W. Lee: *Independent Component Analysis: Theory and Applications*, Kluwer Academic Publishers (1998).

[15] D. MacKay: *Maximum Likelihood and Covariant Algorithms for Independent Components Analysis.* "Draft 3.7" (1996).

[16] M.J. McKeown, T.P. Jung, S. Makeig, G. Brown, S.S. Kindermann, T.-W. Lee and T.J. Sejnowski: *Spatially independent activity patterns in functional magnetic resonance imaging data during the stroop color-naming task.* Proceedings of the National Academy of Sciences USA, 95, pp.803-810,(1998).

[17] J.R. Moeller and S.C. Strother: *A regional covariance approach to the analysis of functional patterns in positron emission tomographic data* J. Cereb. Blood Flow Metab. **11**, A121-A135 (1991).

[18] L. Molgedey & H. Schuster: *Separation of Independent Signals using Time-Delayed Correlations* Physical Review Letters **72** 3634-37 (1994).

[19] K.M. Petersson, T.E. Nichols, J.-B. Poline, and A.P. Holmes. *Statistical limitations in functional neuroimaging I. Non-inferential methods and models.* Phil. Trans. R. Soc. London B **354** 1239-1260 (1999).

[20] M.E. Posner and M. Raichle: *Images of Mind.* Scientific American Library, New York (1994).