Independent Component Analysis in the Presence of Noise in fMRI

D. Cordes[1], R. Nandy[2]

[1]Radiology, University of Washington, Seattle, WA, United States, [2]Psychology, UCLA, Los Angeles, CA, United States

Introduction

ICA is a statistical method for estimating a collection of unobservable signals from observations of their mixtures[1]. An explicit treatment of the noise in fMRI data using ICA is difficult primarily due to the complex structure of the noise. For example, noise studies with phantom data have shown that the Fourier spectrum has low frequency components with approximate 1/f dependence suggesting a low-order autoregressive structure. For better source estimation, it may be necessary to consider additive colored noise in the ICA model. Due to the complexity in dealing with noise, the application of a noisy ICA model has been neglected.

We applied the concept of Gaussian Moments as introduced by Hyvärinen[2,3] to remove the noise-induced asymptotic bias and computed a more accurate mixing matrix in the first stage of ICA. Furthermore, using parametric estimates of the sources densities, we used maximum likelihood estimation to determine the de-noised sources in a second ICA stage.

Theory and Methods

The noisy ICA linear mixing model for mean removed fMRI data can be described as $\boldsymbol{x}_v=\mathbf{A}\boldsymbol{s}_v+\boldsymbol{\eta}_v$ for $v=1,\ldots,N$ where $\boldsymbol{x}_v$ is the observed time course vector at voxel v with T components (time points), $\mathbf{A}$ is the T x p dimensional mixing matrix, $\boldsymbol{s}_v$ is the signal vector (zero mean) with p components at voxel v, $\boldsymbol{\eta}_v$ is the noise vector that follows a multivariate Gaussian distribution with zero mean and covariance matrix $\boldsymbol{\Sigma}$, and N is the number of voxels. To model the noise in fMRI data X, we use an AR(1) model and estimate the parameters σ^2 and ϕ from the tail eigenvalue spectrum of the data covariance matrix. Because of the difficulty of working with the full dataset, we use PCA to reduce the dimension of the input to ICA by $(C-\Sigma)E=E\Lambda$ where $\underset{\sim}{C}=\mathrm{cov}(X)=AA^T+\Sigma$ and E and Λ are the matrices of the major p eigenvectors and corresponding eigenvalues, respectively. The data, $\boldsymbol{X}$, transform according to $\tilde{X}=E^TX=E^TAS+E^T\mathrm{H}$ where $\boldsymbol{S}=[\boldsymbol{s}_1\ldots\boldsymbol{s}_N]$ and $\mathbf{H}=[\boldsymbol{\eta}_1\ldots\boldsymbol{\eta}_N]$ are the matrices of the sources and noises, respectively. The first step in doing ICA is to decorrelate the signals in the data using quasi-whitening leading to an orthonormal mixing matrix of the sources (provided the sources are variance normalized), and a new noise covariance matrix $\boldsymbol{\Xi}$ given by

$$\boldsymbol{\Xi}=(\mathbf{E}^T(\mathbf{C}-\boldsymbol{\Sigma})\mathbf{E})^{-\frac{1}{2}}\mathbf{E}^T\boldsymbol{\Sigma}\mathbf{E}(\mathbf{E}^T(\mathbf{C}-\boldsymbol{\Sigma})\mathbf{E})^{-\frac{1}{2}}.$$

Then, a variant of the FastICA algorithm[2] is used for determining the unmixing vectors $\boldsymbol{w}_i$ (weights):

$$\mathbf{w}_i^*=E[\tilde{\mathbf{x}}\,g(\mathbf{w}_i^T\tilde{\mathbf{x}})]-(\mathbf{I}+\boldsymbol{\Xi})\mathbf{w}_iE[g'(\mathbf{w}_i^T\tilde{\mathbf{x}})] \qquad \text{for } i=1,\ldots,p$$

where w_i* is the updated value of w_i and g is a chosen contrast function approximating the logarithmic derivative of the source densities. In order to de-noise the sources in a second ICA stage, we use maximum likelihood estimation leading to the equation

$$\mathbf{s}_v=\mathbf{B}^T\mathbf{x}_v-\mathbf{B}^T\boldsymbol{\Xi}\,\mathbf{B}\;\;\mathbf{g}(\mathbf{s}_v)$$

In this equation $\mathbf{g}$ is a vector describing the value of the nonlinearities g for each of the p sources $s_v^{(1)}\ldots s_v^{(p)}$ at voxel v. Since in the general case, the source densities cannot be expected to be similar, each source will have a different nonlinearity g. This vector-valued equation is solved iteratively for all realizations of voxels. A complication is provided by the fact that the probability densities of the sources and thus the corresponding nonlinearities g need to be estimated as well. Since data in fMRI are often skew symmetric[4], the nonlinearities g are estimated based on modeling the source probability densities by an asymmetric exponential family of the form

$$p(s)=Ce^{\alpha(s-\beta)+\gamma\sqrt{(s-\beta)^2+|\delta|}}$$ with parameters α, β, γ, δ to be optimized.

Results and Conclusion

Figure 1 shows an example of an asymmetric pdf extracted from real fMRI data and subsequent computation of the nonlinearity using above parameterization. For comparison, the cumulative density function is also shown. Figure 2 shows the accuracy of an estimated source obtained by conventional ICA (blue) and noisy ICA (black) applied to simulated data (30 mixed supergaussian sources from a general exponential distribution with α=0.3 and additive AR(1) noise with ϕ=0.2, 20000 voxels, 300 time frames). Each scatter point corresponds to a voxel. Figure 3 shows the performance of noisy ICA for pseudo-real data as a function of the amplitude c of the added source component. Data used were resting-state data plus a single supergaussian source reflecting a typical fMRI pattern (bilateral motor activation) with strength c was added. The corresponding time course was a Gaussian-convolved boxcar function (periodicity 20 time frames off, 20 time frames on), simulating a typical HRF time course in fMRI. Source accuracy was determined by calculating the correlation coefficient of the extracted source component and original source component (corresponding to the bilateral motor activation). Time course accuracy was determined by calculating the correlation coefficient of the corresponding time courses.

Fig. 1

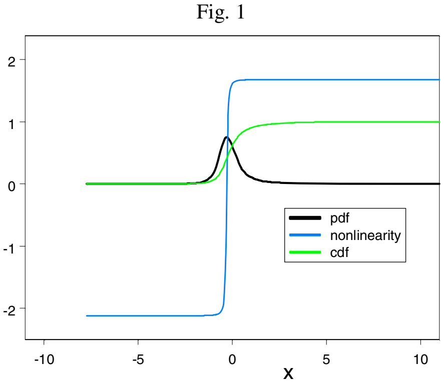


Fig.2

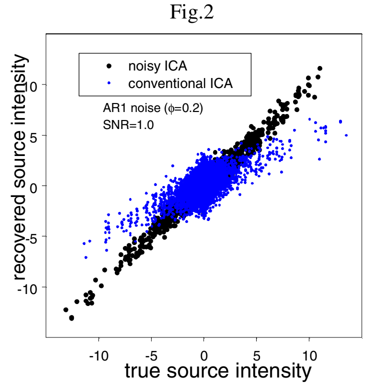
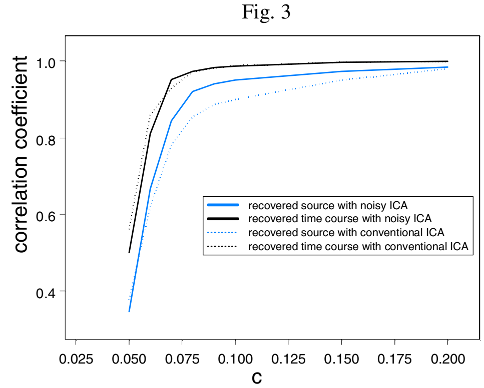


Fig. 3



References

1. McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.P., Kindermann, S.S., Bell, A.J., Sejnowski, T.J. 1998. Hum Brain Map **6**:160-188.
2. Hyvärinen, A. 1999. IEEE Signal Processing Letters **6**(6):145-147.
3. Hyvärinen, A. 1998. Neurocomputing **22**:49-67.
4. Stone, J.V., Porrill, J., Porter, N.R., Wilkinson, I.D. 2002. Neuroimage **15**, 407-421.