COMPARISON OF CANONICAL CORRELATION ANALYSIS AND ICA TECHNIQUES IN FMRI DATA ANALYSIS

Mohammad Reza Arbabshirani [1], Mansoor Nakhkash [1], Hamid Soltanian-Zadeh [2, 3]

[1]Signal Processing Research Lab., Department of Electrical Engineering, Yazd University, Yazd, Iran
[2] Control and Intelligent Processing Center of Excellence, Electrical and Computer Engineering Department, University of Tehran, Tehran, Iran
[3] Radiology Image Analysis Lab., Henry Ford Hospital, Detroit, MI, USA

m.arbabshirani@stu.yazduni.ac.ir, nakhkash@yazduni.ac.ir, hszadeh@ut.ac.ir

ABSTRACT

This paper compares independent component analysis (ICA) and canonical correlation analysis (CCA) for functional magnetic resonance imaging (fMRI) data. The ICA method is implemented with two well-known algorithms, which are Infomax and FastICA. In the CCA method, we have investigated two hemodynamic response models for the signal subspace: differential Gamma and Balloon models. The criterion for the comparison is the area under the receiver operating characteristic (ROC) curve. This criterion is evaluated for different contrast to noise ratios (CNR) for simulated datasets. Also, task-related activation maps are compared for a real auditory dataset. The results indicate the superiority of CCA for CNRs below 0.75; but as CNR goes beyond this limit, the ICA with Infomax algorithm outperforms other methods. Furthermore, the use of differential Gamma and Balloon models provides nearly the same performance.

1. INTRODUCTION

Functional Magnetic Resonance Imaging (fMRI) as an indirect method of brain activity detection, has received great attention in recent years. Each fMRI experiment provides a 4D dataset with low contrast to noise ratio, which makes the analysis a challenging task. Therefore, researchers have proposed several analysis methods that can be categorized into two main groups: hypothesis-based and data-driven.

Hypothesis-based methods, try to utilize prior information about the activation patterns to analyze the fMRI data. The most famous methods of this group are: correlation method, general linear model (GLM) [1] and canonical correlation analysis [2]. Data-driven methods avoid any assumption about the activation pattern and explore the data in order to discover underlying patterns. Principal component analysis (PCA) [3], independent component analysis [4], and cluster analysis [5] are among the data-driven methods.

Canonical correlation analysis (CCA) is a successful extension of the ordinary correlation analysis with a main difference. The correlation method is a univariate technique, whereas the CCA is a multivariate one, i.e. it operates on multidimensional variables. The method combines subspace modeling of the hemodynamic response and the use of spatial dependencies to detect homogeneous maps of brain activity. Comparing with the GLM, the CCA method uses multidimensional datasets on both sides. This makes it more general than the GLM, in which one side is univariate (a voxel time series). The CCA can, therefore, be seen as residing at the top of the hierarchy of model-based methods. Several extensions to the ordinary CCA, such as constrained CCA and nonlinear CCA are also proposed in recent years [6-7].

Among the data-driven techniques, the ICA has been shown to provide a powerful method for the exploratory analysis of the fMRI data [8]. The ICA method assumes linear mixtures of the sources (independent components) and tries to recover the underlying signals based on information theory. Therefore, it is an excellent method to spatially localize and temporaly characterize the sources of blood oxygenation level dependent (BOLD) activation. Both temporal and spatial ICA can be applied to fMRI. Spatial ICA is more popular in fMRI applications because the spatial dimension is much larger than the temporal dimension in fMRI. Among several algorithms developed for ICA, Infomax and FastICA have generated the best results for fMRI data analysis [8-9].

In this paper, a comparison between the CCA and ICA methods is presented. For the CCA method, the same procedure as [7] is followed. Both differential Gamma and Balloon models [10-11] are examined as the hemodynamic response models and spatial filtering is performed using spatial basis functions. As the ICA methods, Infomax [12] and FastICA [13] are applied on the data reduced by the PCA. The comparison is based on the area under ROC curve for different CNRs using simulated datasets and task-related activation maps for a real auditory dataset.

2. CANONICAL CORRELATION ANALYSIS

Canonical correlation analysis is a well known method in statistics, which was first developed by Hotelling in 1936 [14]. For two multidimensional random vectors **x** and **y** that are *n* and *m* dimensional respectively, the CCA seeks linear combinations:

$$X = \boldsymbol{w}_{x_1} \boldsymbol{x}_1 + \boldsymbol{w}_{x_2} \boldsymbol{x}_2 + \ldots + \boldsymbol{w}_{x_n} \boldsymbol{x}_n = \boldsymbol{w}_x^{\,T} \boldsymbol{x} \qquad (1)$$

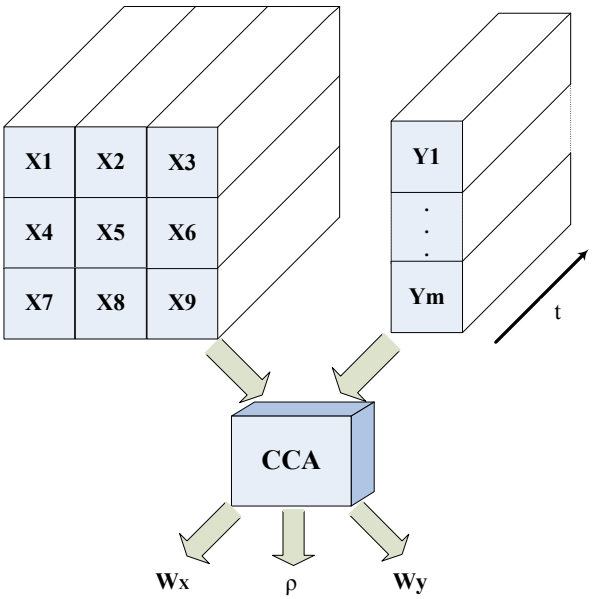


Figure 1. Canonical correlation analysis between a 3×3 region in a fMRI image and a set of *m* basis-functions representative for signal subspace.

$$\boldsymbol{Y} = \boldsymbol{w}_{y_1}\boldsymbol{y}_1 + \boldsymbol{w}_{y_2}\boldsymbol{y}_2 + ... + \boldsymbol{w}_{y_m}\boldsymbol{y}_m = \boldsymbol{w}_y{}^T\boldsymbol{y} \quad (2)$$

so that *X* and *Y* give maximum correlation (which is denoted by ρ). In (1) and (2), the vectors $\boldsymbol{w}_x$ and $\boldsymbol{w}_y$ denote the linear combination coefficients and *X* and *Y* are called canonical variables. A schematic of CCA method is shown in Figure 1.

In fMRI applications, each voxel and its neighbors' time series serve as one of the input sets to the CCA algorithms. We need the other input set, which should be representative of the signal subspace. There are several models for signal subspace. Truncate Fourier series with frequency harmonic of the box car paradigm was proposed in [2]. However, the accuracy of the model is controversial. Later, more accurate signal subspaces based on differential Gamma function and Balloon model were proposed [10-11]. In [7], large numbers of hemodynamic responses were generated by changing the parameters in differential Gamma function and Balloon model within a logical range. Then, the average of the responses and the first principal components served as the signal subspace basis functions. In this paper, we use the same method.

It is seen that CCA, converts simple univariate correlation method to a multivariate method, which incorporates spatial information as well as prior temporal activation pattern. The CCA also conforms well to the connectivity principal in active regions of the brain. It can find a signal in the signal subspace that has similarity to the time course in the region of multiple voxels.

3. INDEPENDENT COMPONENT ANALYSIS

As a blind source separation (BSS) method, the ICA method assumes that the observed signals are the linear mixture of underlying statistically independent source signals. It has recently demonstrated considerable promise in characterizing the fMRI data, primarily due to its intuitive nature and ability for flexible characterization of the brain function. As typically applied, the brain networks are assumed to be spatially nonoverlapping and temporally coherent, though convolutive and other models can be used to relax this assumption. Popular approaches for performing ICA include maximization of information transfer, which is equivalent to maximum likelihood estimation, maximization of non-Gaussianity, mutual information minimization, and tensorial methods. The most commonly used ICA algorithms are Infomax [12], FastICA [13], and joint approximate diagonalization of eigenmatrices (JADE).

Consider an observed m-dimensional data vector $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_m)^T$. The ICA model can be written as:

$$\boldsymbol{x} = A\boldsymbol{s} \quad (3)$$

where $\boldsymbol{s} = (\boldsymbol{s}_1, \boldsymbol{s}_2, ..., \boldsymbol{s}_n)^T$ is an n-dimensional vector, whose elements are independent sources and $A_{n\times m}$ is the unknown mixing matrix. The goal of ICA is to estimate an unmixing matrix $W_{m\times n}$ such that

$$\boldsymbol{y} = W\boldsymbol{x} \quad (4)$$

where $\boldsymbol{y}$ is a good approximation of true sources.

The ICA method is used in fMRI modeling to explore the spatio- temporal structure of the signal, i.e. it can discover either spatial or temporal independent components. Seeking components that are maximally independent in space, spatial ICA (SICA) is more popular for fMRI data analysis. In SICA, the observation data matrix, **X**, is an $N \times M$ matrix (where *N* is the number of time points and *M* is the number of voxels). The aim of fMRI analysis is to factor the data matrix into a product of a set of time courses and a set of spatial patterns. In PCA, this is achieved by the singular value decomposition of the data matrix, where the data matrix is written as the outer product of a set of orthogonal time courses and a set of orthogonal spatial patterns. The ICA is a more general approach since it uses higher order statistical moments. ICA aims at decomposing the data matrix to the product of spatial patterns and corresponding time courses, where either patterns or time courses are independent.

To reduce the dimensionality of the data, before performing ICA, PCA is generally used to discard the principal components with the lowest variance.

4. DATASETS

4.1 Simulated Dataset

In order to make the simulated fMRI data similar to the real data, we superimpose simulated activation time-series on the resting state fMRI data. The resting data sets are acquired using a 1.5T MRI scanner. A total of 256 volumes are collected using a T_2* -weighted gradient echo single-shot echo-planner (EPI) with TR=1.6 sec, TE=50 ms, Flip Angle=90º and FOV=250 × 250 mm². Each volume consists of 15 trans-

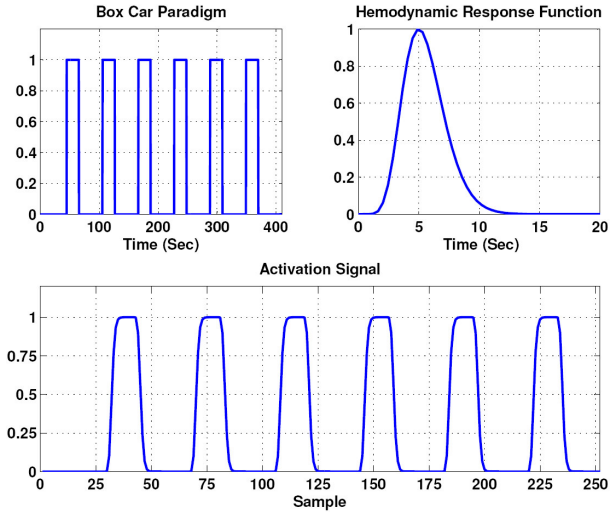


Figure 2. Box Car paradigm, gamma function as HRF and the final activation signal for the simulated dataset.

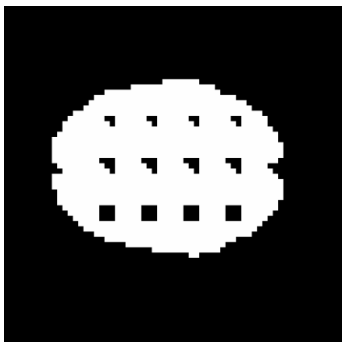


Figure 3. Activation pattern in the simulated dataset

verse slices of size 64 × 64. Activation time series are generated by convolving a boxcar function (6 blocks with 60.8–sec period, 20.8sec on, 40sec off started with 44.8sec rest) with a hemodynamic response function (HRF) that is gamma function with τ and σ set to 5 and 0.06 respectively (Figure 2). Gamma function is defined as:

$$h(t;\tau,\sigma) = e^{-t/\sqrt{\sigma.\tau}} \left(\frac{e.t}{\tau}\right)^{\sqrt{\tau/\sigma}} \quad (5)$$

Simulated activation time-series consisted of 256 points are added to the predefined activation regions (Figure 3) to two consequence slices (slice #9 & #10). Activated regions had different shapes and sizes (6, 12, 18 voxel). The total number of activated voxels was 144 (Figure 3). 7 simulated datasets are constructed with different CNRs ranging from 0.5 to 2.0 by 0.25 increments. Here, the contrast to noise ratio is defined as:

$$CNR = \Delta S / \sigma_n \quad (6)$$

where ΔS is the signal enhancement and σ_n is the noise standard deviation [15].

4.2 Real Audio Task Data

This data set comprises whole brain BOLD/EPI images acquired on a modified 2T Siemens MAGNETOM Vision system. Each acquisition consists of 64 contiguous slices (64x64x64 3mm x 3mm x 3mm voxels). Scan to scan repeat time (TR) was set to 7sec. 96 acquisitions are made from a single subject, in blocks of 6, giving sixteen 42 sec blocks. The condition for successive blocks is alternated between rest and auditory stimulation (6 scans rest, 6 scans stimulation), starting with rest. Auditory stimulation is bi-syllabic words presented binaurally at a rate of 60 per minute. By discarding the first few scans, the initial 96 scans are reduced to 84 scans.

Table 1. Balloon model parameters

Balloon Model		
Parameter	**Base Value**	**Tolerance**
Neuronal efficacy	0.5	0.15
Signal decay	1.2	0.3
Autoregulation	2.4	0.5
Transit time	1	0.5
Capillary transit time	1	0.5
Cp/Cb Ratio	0.01	0.005
Volume Ratio	75	25
Metabolic demand	0.1	0.05
Stiffness	0.3	0.1
Tissue oxygene concentration	0.1	0.05
Tissue Oxygene Scale	5	
Resting oxygene extraction	0.4	0.1

Table 2. Differential Gamma model parameters

Differential Gamma Model		
Parameter	**Base Value**	**Tolerance**
Center Peak 1	5.5	2.5
Shape Peak 1	6	1
Center Peak 2	5	2.5
Shape Peak 2	14.5	2.5
Weight Peak 2	0.125	0.125

4.3 Preprocessing

All fMRI datasets were motion corrected, spatially smoothed and the brain voxels were segmented from the background. For all time-series, mean and drift component were estimated and subtracted from the time-series. For CCA, instead of a single Gaussian smoothing function, a set of 4 spatial basis filters were used to construct a steerable filter as proposed in [7].

5. EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Results of the Simulated Datasets

For CCA, the differential Gamma and Balloon model parameters are summarized in Table 1-2. For each parameter in these models, there is a certain range of values that result in physiologically realistic shapes of the response. By varying the parameters randomly or systematically within these ranges, a large number of plausible response shapes can be produced. In this experiment, 500 HRFs are constructed, and the average of these signals and the first principal component are selected for signal subspace basis.

For ICA, the dimensions of the datasets are reduced by PCA from 256 to 47 using minimum description length

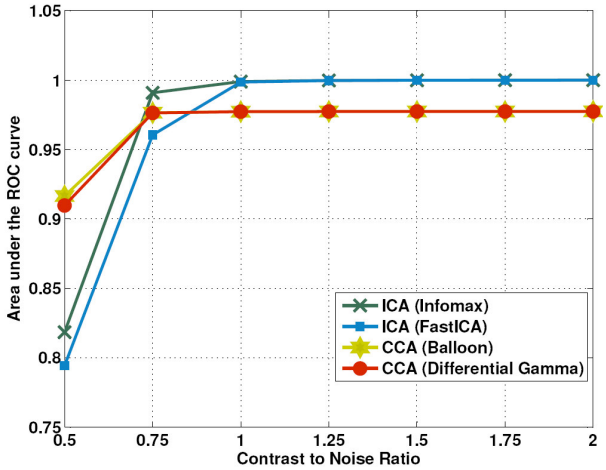


Figure 4. Area under the ROC curve vs. different contrast to noise ratios for the ICA and CCA methods.

Table 3. Area under the ROC curve vs. different contrast to noise ratios for the ICA and CCA methods.

Contrast to Noise Ratio	ICA Infomax	ICA FastICA	CCA - Differential Gamma	CCA Balloon
0.5	0.8185	0.7943	0.9097	**0.9165**
0.75	**0.9908**	0.9605	0.9763	0.9764
1	**0.9989**	0.9985	0.9773	0.9773
1.25	**0.9997**	0.9996	0.9774	0.9774
1.5	**0.9998**	**0.9998**	0.9774	0.9774
1.75	**0.9999**	0.9998	0.9774	0.9774
2	**0.9999**	**0.9999**	0.9774	0.9775

(MDL) criteria [16]. Then, both Infomax and FastICA algorithms are examined for different contrast to noise ratios.

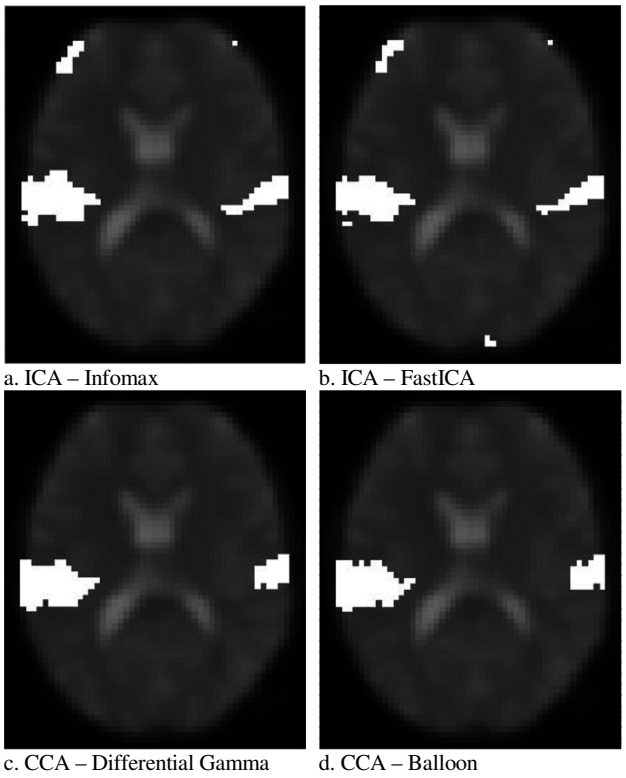
A ROC curve is a plot of 1-β vs. α for different thresholds of the rating scale, where α and β are respectively the probability of type I error (False Positives) and the probability of type II error (False Negatives). In the fMRI context, the rating scale is a parameter reflecting the likelihood of voxel activation. The area under the ROC curve is commonly considered as a good single criterion for characterizing detection accuracy. Thus, the areas under ROC curves for different CNRs are computed for the CCA and ICA methods and the results are summarized in Figure 4 and Table 3.

5.2 Results of the Real Dataset

The CCA and ICA are also applied to the real fMRI datasets. The data reduction for ICA lowers the dimension of the data to 30 by the PCA. Both the correlation-map from the CCA method and the 30 IC-maps from the ICA method are converted to their corresponding z-maps. The z-maps, whose pixels are z-scores, can be determined by

$$Z_i = \frac{X_i - \mu}{\sigma} \qquad (7)$$

For CCA, X_i is the i-th (i ranges from 1 to m×n, the total number of pixels of the map) correlation value of the correlation map, and μ and σ are the mean and the standard deviation of all the correlation values of the same map. For ICA, X_i is the i-th pixel value in the IC-map, and μ and σ are the mean and the standard deviation of all the pixel values of the IC-map.



a. ICA – Infomax b. ICA – FastICA

c. CCA – Differential Gamma d. CCA – Balloon

Figure 5. Activation maps (30th slice) for z-score threshold of 3.1 (p<0.001)

After calculating Z_i, a threshold value of z = 3.1 which is corresponded to the significance level of 0.001 (p<0.001), is applied to z-score maps. Figure 5 shows the resultant activated voxels on the 30th slice.

5.3 Discussion

It seems that for low CNRs, the CCA outperforms the ICA because prior information about the activation pattern helps this method to detect more activated voxels than ICA. By increasing the CNR, the ICA demonstrates a better performance and the CCA results do not change significantly for CNR higher than 1. The performance of Infomax and FastICA algorithms are very similar for CNRs higher than 1, but for lower CNRs infomax algorithm performs better.

It can be concluded that prior information about the activation pattern in CCA, such as signal subspaces based on differential Gamma and Balloon models, bears some inaccuracy. While it helps to detect activation in very low CNR, it limits the performance for high CNRs. It can be resulted that the CCA method still has the disadvantages of hypothesis-based methods. Although HRF and signal subspace models

are improved in recent years, but still there is a considerable gap between the actual brain response patterns and current models. Human brain response varies among different subjects, different brain areas and different times. So, simple linear time-invariant model with experiment paradigm as an input, can not model all these variations and to be very accurate for each person simultaneously.

The ICA method that avoids any prior assumption about the activation patterns, misses some information that is helpful to detect activated voxels in very low CNRs. On the other hand, it can discover the actual brain activation pattern in higher CNRs by exploring the data.

For the real dataset, all the results cover the Brodmann's area (BA) 42 (primary auditory cortex) and BA 22 (auditory associated area). However, CCA detected more continuous activated regions. ICA detects few activated voxels on the other parts of the brain that seems to be false detected non-activated voxels.

Overall, ICA with Infomax algorithm outperforms the other methods investigated in this research in the most of the CNRs range. Also, it should be noted that the computational times of the CCA and ICA methods are comparable.

6. CONCLUSION

In this paper, we compared the CCA and ICA methods for activation detection from the fMRI data. Hemodynamic response subspace is constructed both from differential Gamma and Balloon model for the CCA method. Infomax and FastICA algorithms of ICA with data reduction step based on PCA are also investigated. The results show the overall superiority of ICA method with the Infomax algorithms. However, in very low contrast to noise ratios, CCA performs better. Furthermore, the use of differential Gamma and Balloon models provides nearly the same performance.

ACKNOWLEDGEMENT

The authors would like to acknowledge the Wellcome Department of Imaging Neuroscience at University College, London for providing the auditory fMRI dataset.

REFERENCES

[1] K. J. Friston, A. P. Holmwa, K. J. Worsley, J. P. Poline, C. D. Firth, and R. S. J. Frackowiak, "Statistical parapeter maps in functional imaging: A general linear approach," *Human Brain Mapping*, vol. 2, pp. 189-210, 1995.

[2] Ola Friman, Jonny Cedefamn, Peter Lundberg, Magnus Borga and Hans Knutsson, "Detection of Neural Activity in Functional MRI Using Canonical Correlation Analysis," *Magnetic Resonance in Medicine*, vol. 45, pp. 323–330, 2001.

[3] Lai SH, Fang M., "A novel local PCA-based method for detecting activation signals in fMRI," *Magnetic Resonance Imaging*, vol. 17, pp. 827–836, 1999.

[4] M. J. McKeown, S. Makeig, G. G. Brown, et al. , "Analysis of fMRI data by blind separation into independent spatial components," *Human Brain Mapping*, vol. 6, pp. 160–188, 1998.

[5] M. J. Fadili, S. Ruan, D. Bloyet and B. Mazoyer, "A Multistep Unsupervised Fuzzy Clustering Analysis of fMRI Time Series," *Human Brain Mapping*, vol. 10, pp. 160-178, 2000.

[6] Defeng Wang, Lin Shi, Daniel S. Yeung,"Nonlinear Canonical Correlation Analysis of fMRI Signals Using HDR Models," i*n Proc. IEEE Engineering in Medicine and Biology Conference (EMBC)*, 2005, pp. 5896-5899.

[7] O. Friman., M. Borga, P. Lundberg, and H. Knutsson, "Adaptive Analysis of fMRI Data," *NeuroImage*, vol. 19, pp. 837-845, 2003.

[8] Vince D. Calhoun, Tülay Adali, "Unmixing fMRI with independent component analysis," *IEEE Engineering In Medicine and Biology Magazine*, March/April, pp. 79-90, 2006.

[9] N. Correa, T. Adali, Y. Li, and V.D. Calhoun, "Comparison of blind source separation algorithms for fMRI using a new matlab toolbox: GIFT," *in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Philadelphia, PA, 2005, pp. 401–404.

[10] R. Buxton, E. Wong, and L. Frank, "Dynamics of Blood Flow and Oxygenation Changes During Brain Activation: the Balloon Model," *Magnetic Resonance in Medicine*, vol. 39, no. 6, pp. 855-864, 1998.

[11] K. J. Friston, A. Mechelli, R. Turner, and C. J. Price, "Nonlinear responses in fMRI: the balloon model, Volterra kernels,and other hemodynamics," *NeuroImage*, vol. 12 pp. 466-477, 2000.

[12] A.J. Bell and T.J. Sejnowski, "An information maximisation approach to blind separation and blind deconvolution," *Neural Comput.*, vol. 7, no. 6, pp.1129–1159, 1995.

[13] A. Hyvarinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Comput.*, vol. 9, no. 7, pp. 1483–1492, 1997.

[14] H. Hotelling, "Relations between two sets of variates," *Biometrika,* vol. 28, pp. 321–377, 1936.

[15] N. Lang, "Tutorial in Biostatistics: statistical approaches to human brain mapping by functional magnetic resonance imaging," *Statistics in Medicine*, vol. 15, pp. 389-428, 1997.

[16] Y. Li, T. Adali, and V. D. Calhoun, "Sample Dependence Correction for Order Selection In FMRI Analysis," i*n Proc. ISBI*, Washington, D.C., 2006.