Intrinsic dimension estimation of the fMRI space via sparsity-promoting matrix factorization

Counting the 'brain cores' of the human brain

Harris V. Georgiou
Dept. of Informatics n Telecommunications
National Kapodistrian Univ. of Athens (NKUA/UoA), Greece
xgeorgio@di.uoa.gr

ABSTRACT

Functional Magnetic Resonance Imaging (fMRI) is a powerful non-invasive tool for localizing and analyzing brain activity. This study focuses on one very important aspect of the functional properties of human brain, specifically the estimation of the level of *parallelism* when performing complex cognitive tasks. Using fMRI as the main modality, the human brain activity is investigated through a purely data-driven signal processing and dimensionality analysis approach. Specifically, the fMRI signal is treated as a multi-dimensional data space and its intrinsic 'complexity' is studied via *sparsity-promoting matrix factorization* in the sense of *blind-source separation* (BSS). One simulated and two real fMRI datasets are used in combination with *Independent Component Analysis* (ICA) for estimating the intrinsic (true) dimensionality via detection of statistically independent concurrent signal sources. This analysis provides reliable data-driven experimental evidence on the number of independent active brain processes that run concurrently when visual or visuo-motor tasks are performed. The results prove that, although this number cannot be defined as a hard threshold but rather as a continuous range, however when a specific activation level is defined, an estimated number of concurrent processes or the loose equivalent of 'brain cores' can be detected in human brain activity.

Keywords

fMRI, Independent Component Analysis (ICA), human brain

1. INTRODUCTION

Human brain is the most advanced and efficient signal-processing machine known today. It corresponds to only 2% of the total body weight in adults (about 1.5 kg), yet it consumes 20% of blood oxygen and 25% of glucose, with only 20W at power peak. It consists of roughly 100 billion neurons with 1,000-10,000 synapse interconnections each, packed in 1130-1260 cm^3 of volume, making it the most complex organ in the human body [18, 7, 15]. Analyzing its structure and functionality, especially during the actual process of some cognitive task or in relation to some mental impairment, has been a scientific challenge for centuries. However, only recent technological breakthroughs have enabled the study of the inner workings of living brains. Even today, simulating the structure and only basic neuron functionality of a full-scale human brain in a digital computer is still an infeasible task.

Functional Magnetic Resonance Imaging (fMRI) [16, 15, 19] is a powerful non-invasive tool for localizing and analyzing brain activity. Most commonly it is based on blood oxygenation level-dependent (BOLD) contrast, which translates to detecting localized changes in the hemodynamic flow of oxygenated blood in activated brain areas. This is achieved by exploiting the different magnetic properties of oxygen-saturated versus oxygen-desaturated hemoglobin. In order to properly detect these brain activations and identify the set regions that are relevant to a specific cognitive task, the 3-D space occupied by the brain is partitioned into a grid of 'cubes' or *voxels*. Each voxel constitutes the elementary spatial unit that acts as a signal generator, recorded and registered as a low-resolution 1-D time series. Actual fMRI voxel signals from brain scans can be considered as a mixture of various components or *sources* with different temporal and spatial characteristics. A typical voxel size of 3x3x3.5-5 mm^3 corresponds to roughly 2.5-4 million neurons of several thousands of synapse interconnections each, or $1/40000$ to $1/25000$ of the total brain volume.

This study focuses on one very important aspect of the functional properties of human brain, specifically the estimation of the level of *parallelism* when performing complex cognitive tasks. In some very abstract sense, this is not much different than trying to recover the (minimum) number of actual 'brain cores' required to 'run' all the active cognitive tasks that are registered in the entire 3-D brain volume while performing a typical fMRI experimental protocol that includes visual-only or visuo-motor tasks.

Using fMRI as the main modality, the human brain activity is investigated through a purely data-driven signal processing and dimensionality analysis approach. Specifically, the fMRI signal is treated as a multi-dimensional data space and its intrinsic complexity is studied via *blind-source separation* (BSS) methods. Section 2.1 provides an overview of the fMRI experiments and the nature of sensory data; section 2.2 defines a proper mathematical formulation for

PCI2017 September 28-30 2017, Larisa, Greece

ACM ISBN ....

DOI: ...

the *data unmixing* task and its importance in understanding the true sources of brain activity; section 3.2 provides hints to proper data dimensionality reduction in fMRI; section 3.1 briefly describes ICA as a typical approach for blind-source separation in signal processing; sections 4.1, 4.2.1 and 4.2.2 describe the simulated and real fMRI datasets used in this study; section 5 includes the experiments and results, using all the methods and datasets described earlier; finally, section 6 concludes the study with discussion of the results and their practical meaning.

2. PROBLEM DEFINITION

2.1 The nature of fMRI data

In experimental fMRI procedures, there are two common activation schemes: the *block* paradigms and the *event-related* paradigms [1]. In the block paradigm, the subject is presented with a specific stimulus for a specific time frame, e.g., a set of images of different placement, colors, patterns or categories, and the subject has to press a switch to signal positive or negative feedback as a response. In the event-related paradigm, the subject is exposed to a series of randomized short-time inputs, e.g., a noise or a pain stimulus, with or without the need for specific response from the subject. In both cases, the external input is considered as a primary 'source' and is temporally correlated with the brain activity. Areas of high activation and correlation to the stimulation/response pattern are considered as highly relevant to the specific functional task (visual/motor centers, pain receptors, etc).

The acquired fMRI signal is registered in both spatial (3-D) and temporal (1-D) domain, resulting in a composite 4-D signal. Each spatial axis is registered as a grid of spatial resolution 3-5 mm^3, resulting in a 3-D grid of voxels. Typically, a complete volume of voxel data, e.g. 60x60x30 to 64x64x48, is recorded every 1-2 seconds for a sequence of 100-150 time points [16, 15, 19]. This produces a total of roughly 108K-197K voxels for every time frame or, equivalently, 11-30 million data points organized as a two-dimensional matrix, where each row corresponds to a complete 'snapshot' of brain activity. In practice, the number of actual brain voxels is smaller, since non-brain areas of the grid are masked out before any further processing; however, the data volume still remains within the same order of magnitude. Additionally, typical fMRI experimental protocols involve several subjects, in order to exclude any subject-specific characteristics that may affect the statistical properties of the fMRI data under consideration. Clearly, this creates a high-volume data analysis process that makes it a very complex and computationally demanding task.

2.2 Understanding brain activity

Special matrix factorization algorithms are required to reformulate the fMRI data as a multiplication of two other matrices, where one is for the time courses of the estimated signal 'sources' and one for the corresponding spatial maps of related brain activity. Formally put, if $\boldsymbol{Y} \in \mathbb{R}^{t \times n}$ is the full fMRI data matrix with t rows as time points and n brain voxels 'unwrapped' into a linear vector, then the fMRI data matrix can be factorized as $\boldsymbol{Y} = \boldsymbol{TS}$, $\boldsymbol{T} \in \mathbb{R}^{t \times p}$, $\boldsymbol{S} \in \mathbb{R}^{p \times n}$, where the p spatial maps are collected as rows in $\boldsymbol{S}$ and each column of $\boldsymbol{T}$ contains the activation pattern along time for the corresponding spatial map.

In terms of signal processing, the *hemodynamic response function* (HRF) [16, 15, 19] of the activated neurons, i.e., the changes in oxygen-rich blood flow in the time domain, acts as a low-pass filter in the temporal domain, which in turn modifies the true activation signal that it is registered as fMRI data. The HRF is known to be spatially-varying, which means that there are slightly different hemodynamic responses for different areas of the brain, as well as different HRFs between different brains. Therefore, traditional regression approaches like *General Linear Model* (GLM) approximations [16, 19, 15] that require a pre-defined 'design matrix' are clearly sub-optimal, since this is typically constructed as permutations, transformations, time-shifts and derivatives of one (assumed) 'universal' HRF.

Although the GLM approach is sufficient when only specific sensory-related signal sources (external stimuli) are considered, in the general case it is not possible to define a global design matrix for all signal sources and all (multiple) subjects. Instead, ICA is the most commonly used alternative for this task, in the context of *blind source separation* (BSS) [10, 3, 12] (see section 3.1). In either case, unmixing algorithms are required to be both fast and accurate in identifying the signal 'sources' of fMRI data and the activated areas in the brain corresponding to the specific paradigm source.

3. DIMENSIONALITY ANALYSIS OF THE FMRI DATA SPACE VIA ICA

3.1 Independent Component Analysis (ICA)

In blind source separation (BSS), ICA has been successfully applied to fMRI data for many years [10, 3, 12, 6]. Since the fMRI consists of a mixture of unknown components, corresponding to different brain sources of activity, the unmixing procedure is essentially a BSS problem. However, due to the relatively low temporal and spatial resolution of fMRI data, the non-stationary properties of the signal due to brain- and machine-state variations, as well as the unknown number and exact statistical properties of the sources, the BSS of fMRI data is not a trivial task.

ICA is based on identifying non-Gaussian properties between the sources and separating them from the mixture, essentially reconstructing the original signal as a linear combination of identified components (signal sources), i.e., similarly to the previously discussed formulation $\boldsymbol{Y} = \boldsymbol{TS}$, $\boldsymbol{T} \in \mathbb{R}^{t \times p}$, $\boldsymbol{S} \in \mathbb{R}^{p \times n}$. In this case, $\boldsymbol{S}$ is the matrix of independent components, i.e., spatial maps of brain activity, and $\boldsymbol{T}$ is the mixture matrix, i.e., the corresponding time courses. In fMRI, the ICA can be performed in the spatial or temporal dimension of the (vectorized) voxel data, producing either spatial or temporal components in matrix $\boldsymbol{S}$. Several studies have been conducted in whether spatial or temporal ICA works better for BSS in fMRI data [3]; however spatial maps, i.e., retrieving $\boldsymbol{S}$ as spatial components, seem to be more accurate and useful in most clinical applications of fMRI. The two most common approaches for ICA are the Infomax [2] and fastICA [11, 9, 10] algorithms. Throughout this study, spatial map decomposition is employed via the fastICA algorithm.

In practice, ICA does not include any explicit sparsity-aware constraints, although it is considered sparsity-promoting by enforcing statistical independence between the discovered

components (sources). At the same time it assumes specific statistical properties for the underlying signal sources, i.e, at most one Gaussian distribution and minimal noise artifacts. Hence, ICA unmixing of fMRI data which do not fully satisfy these constraints will construct factorizations that include the maximum allowable number of components for the minimum-error reconstruction of the original (mixed) data. In other words, as described in section 3.2, when the fMRI data include non-trivial mixtures of sources, e.g., as in the case of the simulated dataset (see section 4.1), ICA will construct a factorization model $\boldsymbol{Y} \simeq \boldsymbol{TS}$, $\boldsymbol{T} \in \mathbb{R}^{t \times p}$, $\boldsymbol{S} \in \mathbb{R}^{p \times n}$, with $p = p_{max}$ and non-zero reconstruction error. Similar problems emerge when using explicit sparsity-aware approaches, since they typically produce factorizations with $p \ll p_{max}$ but with larger reconstruction errors, as expected.

3.2 Data decimation and intrinsic dimensionality

One way to deal with the high complexity of the BSS task in fMRI data is to reduce the number of voxels under consideration. Specifically, adjacent neurons in the brain can be considered highly correlated in terms of their responses to external stimuli, provided that the blood vessel networks at very small scales actually introduce some spatio-temporal correlation. If the spatial resolution of the fMRI signal is high, adjacent voxels in the original 2-D or 3-D volume scan can be considered statistically dependent and, hence, redundant. Therefore, some form of decimated voxels set can be used instead as input for the unmixing task, without sacrificing the accuracy of identifying the true inherent sources of the data.

Spatio-temporal correlations between voxels and statistical dependencies are essentially the reason why the fMRI data space has an intrinsic (true) dimensionality much smaller than the number of voxels, i.e., the data matrix $\boldsymbol{Y} \in \mathbb{R}^{t \times n}$ is of column rank $c \ll n$. However, for proper unmixing of the fMRI data, the column rank of matrix $\boldsymbol{Y}$ should be retained even when some decimation process is employed. In other words, the selection of a smaller subset of voxels (instead of all) should be conducted *in a way that does not destroy the information content of the full data*, but instead exploit the the fact that the number of voxels n is very large and their inherent statistical properties can be properly retained with a much smaller subset.

In the cases when only a small set of the signal sources are considered, i.e., the time series of some external stimuli (plus some transformations of it), then regression methods like GLM can be easily formulated with the proper 'design matrix' to recover the related brain activity. When the analysis is conducted in the BSS sense, i.e., all major signal sources are to be recovered (including the stimuli time series), then decomposition methods like ICA provide a well-formulated statistical framework for this task, as long as the proper constraints are asserted as valid (most importantly, the assertion of at most one Gaussian signal source). However, when these statistical assertions are not fully satisfied or when there is a large number of signal sources that are 'exponentially decaying' in terms of importance (contribution to the mixed signal's variance, power spectrum and approximation error), then the number of independent components that ICA or other similar algorithms is limited only by some external pre-defined threshold. In other words, the data matrix $\boldsymbol{Y} \in \mathbb{R}^{t \times n}$ can be factorized only *approximately* as $\boldsymbol{Y} \simeq \boldsymbol{TS}$, $\boldsymbol{T} \in \mathbb{R}^{t \times p}$, $\boldsymbol{S} \in \mathbb{R}^{p \times n}$, with the reconstruction error becoming smaller as the number of recovered components p increases. In theory, if the true sources of the mixed signal are perfectly separable in the BSS sense, then ICA will stop after recovering exactly $p = c$ components, where $c \ll n$ is the column rank of the data matrix $\boldsymbol{Y}$. This means that there are exactly p components, i.e. time courses and corresponding activation maps, that can fully reconstruct the fMRI data for the entire brain activity. Hence, the definition of the optimal value for p by means of non-parametric (data-driven) estimation procedure is of utmost importance in the BSS task for fMRI unmixing.

4. DATASETS

The investigation of fMRI space complexity and intrinsic dimensionality was conducted with two separate types datasets, namely one of simulated fMRI data and two of real fMRI data from carefully designed experiments. The simulated data were introduced as the means to verify the recoverability of the intrinsic dimension when the real signal sources are known and well-defined, while the real data were used as guidelines for estimating the true brain activity in two typical cognitive tasks (visual recognition task and visuo-motor task).

4.1 Simulated fMRI datasets

In this study, an adapted version [14] of the real-valued fMRI data generator code from the MSLP-Lab [17] toolbox was used for creating artificial fMRI data as a mixture of eight main sources [6]. Using the basic knowledge of the underlying statistical characteristics of the underlying sources, the components include three highly super-Gaussian sources (S1, S2, S5), a Gaussian source (S4) and a sub-Gaussian source (S3), plus two more super-Gaussian sources (S6, S8) and a sub-Gaussian source (S7). The time course for each component defines the temporal characteristics of the corresponding source, namely one task-related (S1), two transiently task-related (S2, S6) and several artifact types (S3, S4, S5, S7, S8), including respiration, cardiac pulsation, scanner drift, background noise, etc. These sources can be considered as spatial maps that are activated according to their time course and mixed linearly to produce the final (simulated) fMRI data.

Although in typical fMRI experiments there is only one sensory 'input' (stimulation), here the full set of eight sources (S1...S8) was considered throughout the evaluation. Specifically, the simulated fMRI data included eight spatial maps of size 60x60 voxels (2-D 'slices') and a 100-point time course, with statistical properties as described above. Each spatial map was linearized by row-concatenation into a (row) vector of 3600 voxels, registered along its time course (column) vector of 100 points. Finally, these eight 100x3600 matrices of spatio-temporal maps were mixed linearly to produce the final eight-source mixing of simulated fMRI data into one matrix of that same size. Hence, in terms of the problem formulation presented in section 2.2, the final matrix of (simulated) fMRI data is registered as $\boldsymbol{Y} \in \mathbb{R}^{t \times n}$, where $t = 100$ time points and $n = 60^2 = 3600$ voxels. Since the final data matrices are always linearized in a similar way before applying any unmixing algorithm like GLM or ICA, using 2-D 'slices' of (simulated) voxels instead of full 3-D (real) brain scans in each time point affects only the volume of the data and not the task itself.

4.2 Real fMRI datasets

4.2.1 *ds101 – The 'Simon' task*

The 'NYU Simon Task' dataset [13] comprises of data collected from 21 healthy adults while they performed a rapid event-related Simon task. In this study, the data from nine (out of 21) low-noise subjects were used, including two runs each, for a total of 18 fMRI scans. Each dataset was masked for exclusion of non-brain areas and subsequently thresholded for exclusion of brain areas with near-zero activity, in order to suppress any noisy artifacts. The resulting number of voxels ranged roughly between 28K and 39K, while the number of snapshots was fixed to 151 time points. In terms of the formulation of section 2.2, each fMRI data matrix is $\boldsymbol{Y} \in \mathbb{R}^{t \times n}$ with $t = 151$ time points and $27631 \leq n \leq 38735$ 'non-zero' voxels.

Three variants of each dataset were used, regarding the smoothing pre-filtering. Specifically, according to standard fMRI acquisition practice, a Gaussian smoothing kernel was applied to the original 3-D voxel space, in order to suppress noise artifacts and enhance the spatial continuity of the voxel data. With respect to their *Full Width at Half Maximum* (FWHM) [4, 5], or $2\sqrt{2 \cdot ln2} \cdot \sigma \simeq 2.35482 \cdot \sigma$ for Gaussian kernels, two different spatial sizes were used: 4 mm^3 and 8 mm^3. In practice, since the voxel resolution in this dataset is 3x3x4 mm^3, the smaller kernel performs (softer) averaging on 1-1.33 neighboring voxels, while the larger kernel performs (more aggressive) averaging on 2-2.67 neighboring voxels. These two 'smoothed' versions, plus the original non-smoothed version, are the three variants of each fMRI scan, used throughout the experiments (see section 5 for details).

4.2.2 *ds105 – Visual object recognition task*

The 'Visual Object Recognition Task' dataset [8] comprises of data collected from six healthy adults while they performed a visual recognition task. The selection, preprocessing and smoothing stages performed here are the same as for the 'ds101' dataset. The data from six (all) subjects were used, including three (out of 12) runs each, for a total of 18 datasets of fMRI scans. Each dataset was masked for exclusion of non-brain areas and subsequently thresholded for exclusion of brain areas with near-zero activity. The resulting number of voxels ranged roughly between 22K and 47K, while the number of snapshots was fixed to 121 time points. In terms of the formulation of section 2.2, each fMRI data matrix is $\boldsymbol{Y} \in \mathbb{R}^{t \times n}$ with $t = 121$ time points and $22387 \leq n < 47192$ 'non-zero' voxels. Three variants of each dataset were used, i.e., the original and two smoothed versions of each fMRI scan, described above.

5. EXPERIMENTS AND RESULTS

5.1 ICA analysis of the datasets

The ICA experiments that were conducted with the simulated fMRI data included two distinct realizations of the dataset, generated by the same procedure and the same specifications as described in section 4.1. Since the data generation includes several noise components, the two realizations were used as an additional verification check to validate that slightly different mixtures of (artificial) fMRI data do not produce significant differences in the ICA error-versus-components plots.

Figure 1: Ideal (blue) and ICA-recovered (red) time courses of the eight sources in the simulated fMRI dataset.

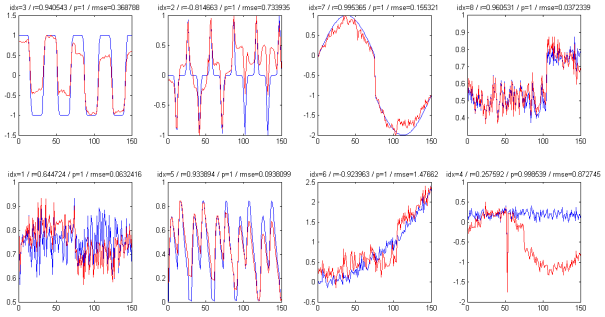


Figure 2: ICA-recovered activation maps of the eight sources in the simulated fMRI dataset.

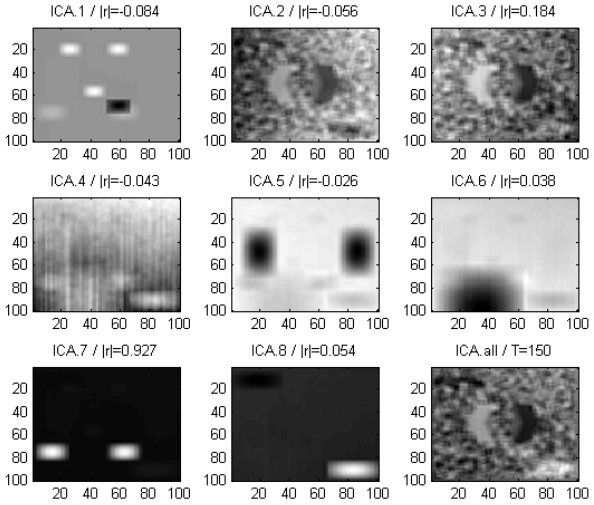
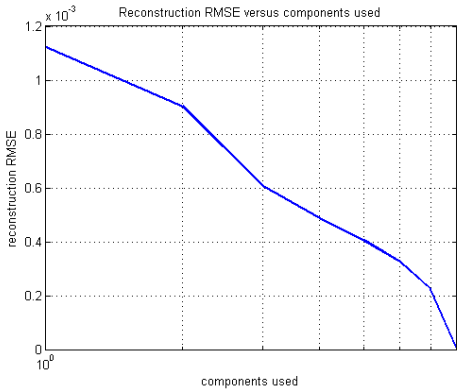


Figure 1 presents the time courses of the ICA factorization (matrix $\boldsymbol{T}$), with the blue curves representing each of the eight ideal (true) sources and the red curves representing the corresponding ICA-recovered sources. Ideal (blue) and ICA-recovered (red) time courses of the eight sources in the simulated fMRI dataset are illustrated. Parameter r is the correlation coefficient between the original (ideal) and the recovered time course, p is the corresponding p-value and *rmse* is the matching error. The first component (upper-left corner) corresponds to the pre-defined external stimuli.

Figure 2 presents the corresponding activation maps (matrix $\boldsymbol{S}$) recovered by ICA, spatially reshaped into proper 2-D brain 'slices', where the reconstruction errors are visible as artifacts ('ghost' artifacts). ICA-recovered activation maps of the eight sources in the simulated fMRI dataset are illustrated, spatially reshaped into proper 2-D brain 'slices'. The lower-left box corresponds to the activation areas for the pre-defined external stimuli. The lower-right box illustrates the complete reconstructed fMRI mixture at the final time point $t = 150$.

Figure 3 presents the plot of reconstruction error (RMSE) versus the number of ICA components used. Specifically, after the ICA unmixing model is complete, the ICA components are used one by one in rank-1 reconstructions of

Figure 3: Reconstruction error versus number of used components.



the original data and the corresponding errors are used for sorting the components in ascending order (smallest RMSE first). Subsequently, a set of components starts from the first one (top of the list) and increased by adding the next one in each step, while estimating and registering the corresponding reconstruction error. The plot illustrates the total reconstruction error decreasing almost exponential-linearly as the number of used components increases, as expected. It should be noted that for 'perfect' ICA factorizations, as in the case of simulated fMRI data, the number of components recovered by ICA is exactly the same as the number of signal sources (true) used in the mixture that created these data (see section 4.1).

The ICA experiments that were conducted with the real fMRI data included two distinct datasets, 'ds101' and 'ds105', as described in sections 4.2.1 and 4.2.2, respectively. Instead of a single 2-D brain 'slice' as in the case of the simulated fMRI data, here the datasets employ full 4-D fMRI data, i.e., 3-D voxel grid of the brain volume evolving in 1-D time course.

Figure 4 illustrates the top-10 of the 50 ICA-recovered time courses of components in a sample run with the 'ds101' dataset. Although the ICA converged successfully with the minimum attainable reconstruction error, the unmixing model failed to detect one single component that closely matches the ideal time course of the stimuli. However, it is evident that one component (third from top-left) matches component no.7 and two components (upper/lower left) match component no.8 of the simulated fMRI data as illustrated in Figure 1 in terms of overall shape and noise characteristics.

With respect to reconstruction error versus number of used components, Figure 5 and Figure 6 illustrate how the RMSE changes (drops) as the size of the ICA mixture becomes larger. Red curves represent the RMSE against the number of used components up to an upper limit of 10, 25, 50 and 100. The final (right-most) point in blue represents the maximum-size, lowest-RMSE in each case. Hence, the general slope of the red curves, as well as the dotted blue line connecting the end points, illustrate the robustness of the ICA unmixing process in each of the real fMRI datasets.

6. DISCUSSION

This study presents a purely data-driven approach to the estimation of the level of *parallelism* in human brain. In

Figure 4: 'ds101' (non-smoothed), top-10 of the 50 ICA-recovered time courses of components in a sample run.

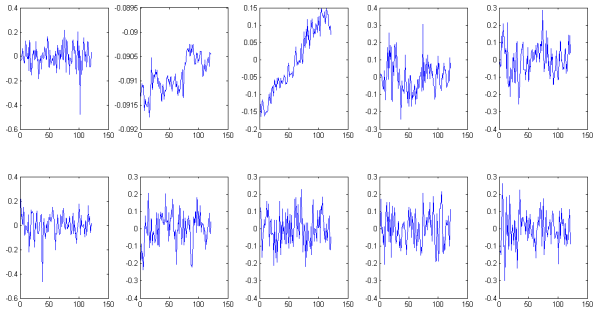


Figure 5: 'ds101' (non-smoothed), ICA reconstruction error versus number of used components.

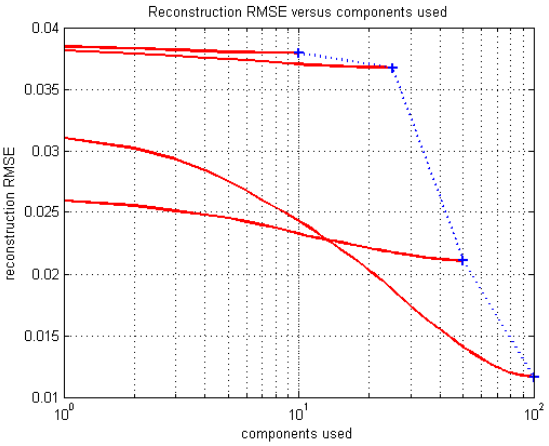
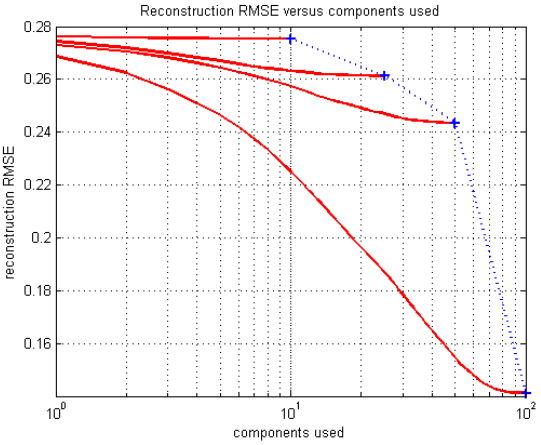


Figure 6: 'ds105' (non-smoothed), ICA reconstruction error versus number of used components.



some very abstract sense, the experimental protocol described here is not much different than trying to recover the (minimum) number of corresponding 'brain cores' required to 'run' concurrently all the active cognitive tasks that are registered in the entire 3-D brain volume while performing visuo-motor cognitive tasks.

The results show that ICA can indeed address the unmixing task with moderate to good performance, especially with regard to the signal sources related to well-defined external stimuli (see component no.7 in Figure 1 and Figure 2). For the simulated fMRI dataset, the total RMSE for the ICA mixture, reconstructing the original signal with all the recovered (eight) components, is practically zero (see Figure 3). The most important results in this case are: (a) the number of ICA components recovered matches the number of true sources used to construct the original mixture and (b) one of the recovered components closely matches (highly correlated) with the well-defined external stimuli (square-shaped time course). This is extremely important in real fMRI experimental protocols, where specific stimuli types are to be correlated to specific brain areas for constructing 'global' brain *atlases*.

As described in sections 3.1 and 3.2, in the case of real fMRI datasets the ICA factorization is only approximate (RMSE is never zero) and the minimum reconstruction error is achieved only when using the maximum allowable number of components. On the other hand, from Figure 5 and Figure 6 it is clear that the reconstruction error rises sharply when the number of used components is much lower than this upper limit. For the 'ds101' dataset, this number seems to be somewhere in $25 < p < 50$ (see Figure 5), while for the 'ds105' dataset it is $p \simeq 50$ (see Figure 6). In both cases, the non-smoothed variants of the datasets were used, hence there is no loss of fine-detail activations and these estimations can be considered as realistic and consistent.

In short, analysis of the non-smoothed variants of the real fMRI datasets (i.e., no information loss) proved that even when performing complex visuo-motor tasks, the number of independent brain processes are in the order of 50. This means that, in theory, an artificial equivalent of a brain-like cognitive structure may not require a massively parallel architecture at the level of single neurons, but rather *a properly designed set of limited processes that run concurrently on a much lower scale*. Hence, although current state-of-the-art VLSI technologies still include very limited features and processing power when compared to the real human brain, the assertion of employing actual parallelism level of much lower order can provide useful hints to future projects.

Acknowledgments

The author wishes to thank professor Sergios Theodoridis, Yiannis Kopsinis and Anastassios Fytsilis, colleagues at the Dept. of Informatics & Telecommunications, National Kapodistrian Univ. of Athens (NKUA/UoA), for their collaboration in related projects and the useful discussions for works-in-progress in fMRI analysis and sparse modeling.

7. REFERENCES

[1] M. Behroozi, M. R. Daliri, and H. Boyaci. Statistical analysis methods for the fmri data. *Basic and Clinical Neuroscience*, 2(4):67–74, 2011.

[2] A. J. Bell and T.J. Sejnowski. An information maximisation approach to blind separation and blind deconvolution. *Neural Computation*, 7:1129–1159, 1995.

[3] V. D. Calhoun, T. Adali, G. D. Pearlson, and J. J. Pekar. Spatial and temporal independent component analysis of functional mri data containing a pair of task-related waveforms. *Human Brain Mapping*, 13(1):43–53, May 2001.

[4] M. K. Chung. Gaussian kernel smoothing. *Lecture notes*, pages 1–10, 2012.

[5] M. K. Chung. *Statistical and Computational Methods in Brain Image Analysis*. CRC Press, USA, 2013.

[6] N. Correa, T. Adali, Yi-Ou Li, and V.D. Calhoun. Comparison of blind source separation algorithms for fmri using a new matlab toolbox: Gift. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, volume 5, pages v/401–v/404 Vol. 5, March 2005.

[7] K. P. Cosgrove, C. M. Mazure, and J. K. Staley. Evolving knowledge of sex differences in brain structure, function, and chemistry. *Biol Psychiat*, 62(8):847–855, 2007.

[8] J. Haxby, M. Gobbini, M. Furey, A. Ishai, J. Schouten, and P. Pietrini.

[9] A. Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. *Neural Networks*, 10:626–634, 1999.

[10] A. Hyvarinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. Wiley-Interscience, 2001.

[11] A. Hyvärinen and E. Oja. A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 9(7):1483–1492, 1997.

[12] Andreas Jung. An introduction to a new data analysis tool: Independent component analysis. In *Proceedings of Workshop GK" Nonlinearity"-Regensburg*, 2001.

[13] A. M. Kelly and Milham M. P. ds101: Nyu simon task – openfmri.org (dataset), 2011.

[14] Yannis Kopsinis, Harris Georgiou, and Sergios Theodoridis. fmri unmixing via properly adjusted dictionary learning. In *22th European Signal Processing Conference (EUSIPCO 2014)*.

[15] N. Lazar. *The Statistical Analysis of Functional MRI Data*. Springer, 2008.

[16] Martin A. Lindquist. The statistical analysis of fMRI data. *Statistical Science*, 23(4):439–464, November 2008.

[17] MLSP-Lab. Machine learning for signal processing laboratory – http://mlsp.umbc.edu/.

[18] A. Parent and M. B. Carpenter. *Carpenter's Human Neuroanatomy*. Williams-Wilkins, 1995.

[19] S. M. Smith. Overview of fmri analysis. *The British Journal of Radiology*, 77:S167–S175, 2004.