

Brains and Phantoms: An ICA Study of fMRI

Jarkko Ylipaavalniemi¹, Seppo Mattila^{2,3}, Antti Tarkiainen³,
and Ricardo Vigário¹

¹ Neural Networks Research Centre,
Laboratory of Computer and Information Science,
Helsinki University of Technology, P.O. Box 5400, FI-02015 TKK, Finland
jarkko.ylipaavalniemi@tkk.fi

² Brain Research Unit, Low Temperature Laboratory,
Helsinki University of Technology, P.O. Box 2200, FI-02015 TKK, Finland

³ Advanced Magnetic Imaging Centre,
Helsinki University of Technology, P.O. Box 3000, FI-02015 TKK, Finland

Abstract. Biomedical signal processing is arguably the most successful application of independent component analysis (ICA) to real world data. For almost a decade, its use in connection with functional magnetic resonance imaging (fMRI) has allowed for data-driven analysis, partly removing the constraints for stringent experimental setups, which are often required by traditional methods based on the use of temporal references. Recent studies on the consistency of independent components have resulted in a series of tools enabling a more reliable use of ICA. In particular, it is now rather easy to detect algorithmic overfitting and isolate subspaces of related activation. Yet, often the nature of the components may not be determined unambiguously. Focal fMRI signals, seemingly originating from within a subject's brain and showing physiologically plausible temporal behavior, are typically considered relevant. This paper presents a study, which makes use of a standard homogeneous spherical phantom and shows evidence for artifacts caused by the measuring device or environment, with characteristics that could easily be misinterpreted as physiological. Our results suggest that reliable analysis of fMRI data using ICA may be far more difficult than previously thought. At least, artificial behavior revealed by phantom analysis should be considered when conclusions are drawn from real subject measurements.

1 Introduction

Searching for a set of generative source signals from their linear mixing, with little to no knowledge on the sources or the mixing process, is referred to as blind source separation (BSS). Independent component analysis (ICA) is possibly the most widely used data-driven method to solve such problems (a good introduction to ICA, including its historical debuts and theoretical frameworks can be found in the textbook [1]; further reading and applications can also be found in [2, 3]). Biomedical signal processing is arguably the most successful application of ICA to real world data (representative examples can be found in, *e.g.*, the following review papers [4, 5, 6, 7]).

On the other hand, functional magnetic resonance imaging (fMRI) has secured a strong position in non-invasive studies of the living human brain. It provides indirect information on neural activity, by measuring the blood oxygenation level dependent (BOLD) signal (*cf.*, [8]). When analyzing fMRI data under the statistical parametric mapping (SPM) framework [9], researchers validate an active brain region through the match of its temporal activation pattern with a carefully predetermined experimental setup. Therefore, they can only validate predictions. When performing a data-driven analysis, such as ICA, the researcher is given greater freedom and is thus capable of detecting unforeseen activity. This allows the study of a whole new set of more complex research questions. Yet, the problem of interpreting the nature of the detected components remains, since not all components have a physiological origin. The key rationales often used in identifying components of interest include the focal nature and potential symmetry of the spatial patterns; whether the activation is located inside the brain and if it falls on expected regions for given stimuli; and how plausible the corresponding time-courses are.

Recent studies on the consistency of independent components resulted in a series of tools enabling a more reliable use of ICA (*cf.*, [10, 11, 12]). In particular, it is now rather easy to detect algorithmic overfitting and isolate subspaces of related activation. However, some results (*cf.*, [13, 11, 14]) suggest that identifying the relevant components may, in fact, not be as straightforward as previously thought, *e.g.*, in the presence of artifacts with characteristics matching the aforementioned rationales on activation volumes and temporal patterns.

In this paper we confront the analyses of brain responses to auditory stimuli (presented earlier in [14]) and of recently collected data from a standard MRI phantom. Our results suggest that the analysis of fMRI data using ICA may be more difficult than previously thought.

2 Data and Methods

2.1 Functional Magnetic Resonance Imaging

Varying concentrations of oxygen in the blood result in changes of its magnetic properties. Because active brain areas produce a local increase in the blood flow, measuring the MR-signal during rest periods and during task conditions, *e.g.* when attending to stimulus presentation, results in detectable differences in the measured images. This is the general basis of fMRI.

When using controlled stimuli, it is common to look for voxels in the brain with a temporal activation pattern that matches the time-courses of the stimuli (*cf.*, [9]). However, the use of general data-driven methods, such as ICA, have been suggested when attempting to observe epiphenomena that are hard to tie to the stimuli or tasks, or when searching for brain reaction to unlabeled stimuli.

2.2 Phantom fMRI Measurements

Artifacts caused by the measuring device or environment can corrupt the fMRI signal. Hence, their characteristics have to be assessed in order to reliably sepa-

rate them from genuine BOLD signals. One example of a known artifact in fMRI signal is the low frequency drift [15], but also more complex artifacts do exist.

Test objects, *i.e.*, phantoms, are often used for assessing the quality of data collected by MRI equipments (*cf.*, [16, 17]), in terms of image properties, such as, signal-to-noise ratio and image uniformity. They can also be used for testing and guiding further analysis applied to the data (*cf.*, [13]).

Here, we use a homogeneous spherical phantom, provided by GE Healthcare. It has a diameter of about 15 cm, and is filled with silicone gel, in order to produce images with intrinsically uniform brightness over the whole phantom. Half an hour before starting the measurements, the phantom was placed within the head coil and the patient bed moved into measurement position to avoid any movement induced artifacts. We used the Gradient Echo (GRE) Echo Planar Imaging (EPI) technique, commonly used for fMRI. All the data were acquired using a 3.0 Tesla MRI scanner (Signa EXCITE 3.0T; GE Healthcare, Chalfont St. Giles, UK), with a quadrature birdcage head coil at the Advanced Magnetic Imaging Centre of Helsinki University of Technology.

The data consisted of 300 fMRI time points with $TR = 3000$ ms (*i.e.*, a 3 second time resolution). The first four scans were excluded from further processing. For each time point, we acquired 37 axial 3.0 mm slices (spacing = 0 mm) with a 96×96 acquisition matrix and a field of view of 200 mm. We used a flip angle of 90 deg and $TE = 32$ ms, all typical values in fMRI studies. Since we were interested in seeing possible artifacts arising from the measurement equipment or the environment, the data was not preprocessed before the analysis.

2.3 Real Auditory Measurements

To compare the components found from the phantom data to the ones found from real measurements, we also used fMRI data of 14 human subjects attending auditory word stimuli. The stimuli consisted of repetitions of resting and listening periods. The data consisted of 80 fMRI time points and were acquired with the same imaging parameters, head coil, and MRI scanner as the phantom data (further information about the experimental paradigm, and data analysis with a reliable ICA procedure can be found in [14]). However, in contrast to the phantom data, these measurements were acquired prior to the EXCITE upgrade for the imaging equipment.

2.4 Reliable ICA

In BSS, the measured data is an instantaneous linear mixture of generative source signals, *i.e.*, $\mathbf{X} = \mathbf{AS}$, for \mathbf{X} , \mathbf{A} and \mathbf{S} , respectively, the observed data, the mixing matrix and the underlying sources. The goal is to identify both the sources and the mixing process with as few assumptions as possible. ICA solves the BSS problem by assuming only that the generative sources are statistically independent from each other. Hence, when applied to fMRI data, we often look for spatially independent neuronal activity, with the columns of the mixing matrix giving the temporal activation of such components.

Theoretically, statistical independence means that the joint probability density of the sources is factorisable on its marginal densities. In practice, several estimation algorithms have been proposed to perform ICA, mainly based on concepts such as negentropy, mutual information or maximum likelihood (for further information, *cf.*, [1, 2, 3]). The experiments in this study use the FastICA algorithm [18], an iterative method with fixed-point optimization. The considerations based on the FastICA algorithm should also be valid for other implementations.

When analyzing a finite data set, the estimated components may change slightly each time the analysis is performed. This behavior can be caused by many factors. For example, the theoretical assumption of statistical independence may not hold for the analyzed data [19, 4]. In this case, the somewhat less restrictive sparse constraint for the underlying sources may still hold, as suggested in [5]. Also, the algorithmic implementation of ICA may be inherently stochastic. Furthermore, additive noise or other data features can cause variations in the solutions. When the degrees of freedom is high, there is also a tendency for overfitting the data [20]. For ICA, this corresponds to bumps or spikes, which occur quite randomly each time the algorithm is run.

In this paper, the consistency of the estimated signals is tested by running the algorithm with many different initial conditions, and bootstrapping of the data. FastICA was used in *symmetric* mode with *tanh* nonlinearity, other parameters where left at default values. The solutions found are combined according to their similarities. Estimates that differ greatly from run to run are less likely to correspond to true components, whereas the ones with small variances are considered reliable (further details, including the combination strategy, can be found in [14]). Similar approaches for consistency analysis and visualization can be found in, *e.g.*, [10, 11, 12].

3 Results

A set of independent components showing the simplest or the most structured time-courses found by analyzing the phantom measurement is shown in Fig. 1. The disks on the left hand side show the spread of the estimates from the different runs. The upper disk depicts the intra-group and the lower disk the inter-group distances. The circles represent mean distances. In ideal estimate discrimination, the upper disk should fit within the hole in the lower disk. The slices represent, clockwise from top-left, the sum of the whole volume; the slice containing the highest power; the slice with the maximum voxel value; and the slice with the minimum voxel value. The mean time-course is shown as a line, superimposed on the spread from all the runs, shown as quantile bands with different intensities. More details on the reliability measures can be found in [14].

All the shown components were found to be very reliable, although some exhibit a small amount of variability. It appears that none of the components results from the signal processing related to the analysis itself, or random noise in the data. Furthermore, the time-courses of the components show clear and

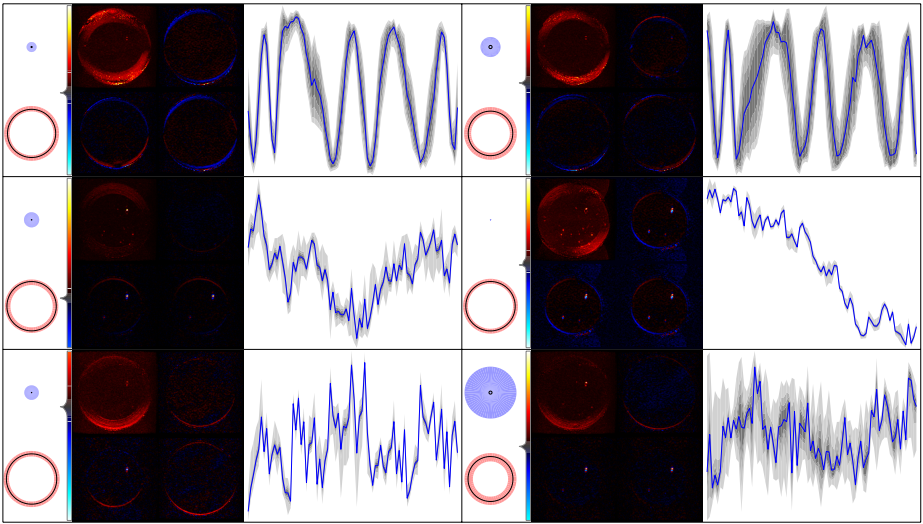


Fig. 1. A set of independent components with simple or structured time-courses, found from the phantom measurement. A time period of 240 seconds is shown.

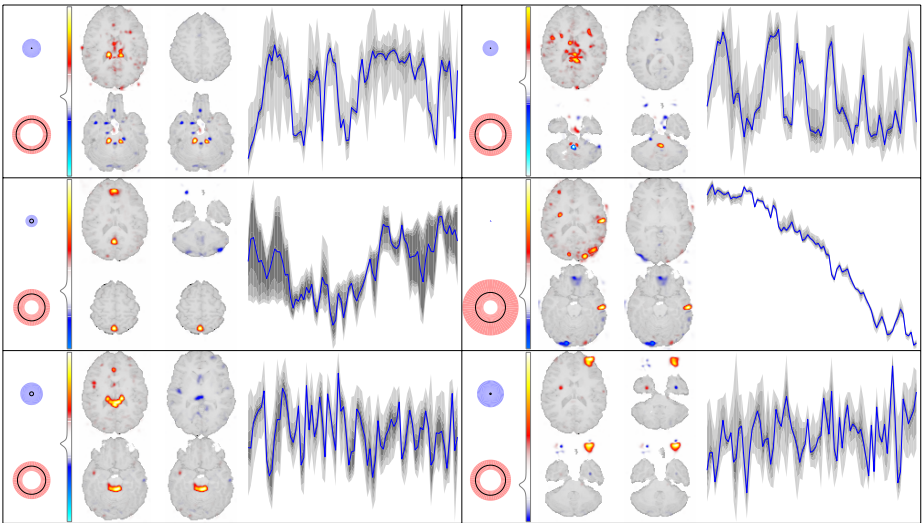


Fig. 2. A set of independent components, found from the measurement of a real subject, which resemble the ones found from the phantom data. A time period of 240 seconds is shown.

systematic structures on time scales of several seconds. In other words, components like these could be considered as relevant activity, in real brain data.

Another set of components found from the real brain measurements with time-courses resembling the ones found from the phantom data are shown in Fig. 2. The results are from a single subject, but similar components were found in the data of all the 14 subjects studied. Here the slices are superimposed on a structural MRI of the subject. The first component has a well structured time-course, which could easily be misinterpreted as being related to the on-off type stimuli. In addition, the volume contains a very focal and symmetric activation pattern in the mid-brain area. However, the time-course, *i.e.*, mixing vector, shares so much of the characteristics of the first component of the phantom data, that a more probable explanation is that both components are manifestations of the same scanner or environment induced artifact. This, however, would be impossible to notice without having performed the phantom analysis.

Similarly, the other components shown in Fig. 2 exhibit focal activation patterns within the brain and/or structured time-courses, *e.g.*, periodic or slowly varying. Therefore fitting the aforementioned rationales for identifying interesting components. A close match for many of the time-courses can be found from the phantom data, with correlations reaching as high as 90%. However, simple mathematical measures, such as cross-correlation do not show the complete picture. For example, temporal delays may need to be taken into account. Also, at times, the phantom components show temporal behavior close to the one used in the block design of the experiment, causing very strong correlations with several components. The crucial question is whether in human evaluation the characteristics of the phantom and brain components are confusingly similar, *e.g.* the time-course associated with the primary auditory cortex is remarkably close to the first two phantom components in Fig. 1 (mean correlations 43% and 38%, respectively). However, in this study, the auditory component can unquestionably be labeled as physiological. Yet, it may still contain contribution from an artificial signal. Furthermore, artificial signals can be much harder to rule out in less controlled experimental setups.

The spatial differences between the phantom and brain components may be attributed to the homogeneity of the phantom, in contrast to the highly non-uniform MR signal of the human brain. The structural differences may also affect the magnitude of the measured components. Some of the components could be related to, *e.g.*, heating of the gradient coils during the imaging or time-dependent changes in the magnetic fields. Other hardware instabilities and the imaging environment can also produce artifacts (*cf.*, [15] and references therein). However, a detailed discussion on the origin of the components is beyond the scope of this paper.

With a real subject, artifacts can also be caused by, *e.g.*, cardiac pulsation or head movements. For instance, some of the components in Fig. 2 show characteristics typical for head movements that have not been completely compensated in the preprocessing of the data. Interestingly, similar behavior in the phantom data suggests that they may in fact be caused by other phenomena.

4 Discussion

Phantom measurements are routinely used for verifying and calibrating the quality of MRI machinery. However, data-driven analysis of phantom fMRI data has been largely overlooked, possibly due to the lack of a method for assessing the reliability of the solutions. Although some earlier work (*cf.*, [13]) has shown that consistent independent components can indeed be found from phantom measurements, to our knowledge, such components have never before been shown publicly and compared with components found from fMRI studies with real subjects. However, the results presented here strongly suggest that such comparisons may be very valuable for the whole research field.

The presented results from analyzing phantom data using ICA reveal evidence for possible misinterpretations in ICA studies with real subjects. The evidence suggests that analyzing fMRI data using ICA may actually be far more difficult than previously thought. It is possible that other methods than ICA are also affected. For example, the reference time-course of the stimuli could have points of coincidence with artificial signals. The analysis would thus mix real brain activations with artifacts.

Although not shown, the results also suggest that the imaging parameters affect the scanner induced components. Therefore, it is important that the phantom measurements are made with the same parameters as those used with the real subjects. It is also expectable that the artifacts can differ with, *e.g.*, time, scanner and measurement coils used. This suggests that data-driven analysis, such as ICA, of phantom data may be useful for quality control of fMRI machinery. The possible effects of different preprocessing steps, typical in fMRI analysis, could also be tested with a similar approach.

The purpose of this paper is to be a word of warning for the ICA community involved in analyzing fMRI data. Clearly, we need a better understanding of the artificial, scanner or environment induced, signals, and of the way they are manifested in phantom and real brain measurements. Possible methods for automatic exclusion of such artifacts should also be considered. If artifacts with systematic characteristics are observed, they could be used for designing real brain measurements such that the stimulus timing does not coincide with the known artifacts. However, the present results strongly suggest that if a researcher wants to base conclusions on components with a purely physiological origin, the ICA results should be compared with phantom measurements.

References

1. Hyvärinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. 1st edn. Wiley-Interscience, New York, NY (2001)
2. Cichocki, A., Amari, S.I.: Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. 1st edn. Wiley-Interscience, New York, NY (2002)
3. Stone, J.V.: Independent Component Analysis : A Tutorial Introduction. 1st edn. MIT Press/Bradford Books, Cambridge, MA (2004)

4. Jung, T.P., Makeig, S., McKeown, M.J., Bell, A.J., Lee, T.W., Sejnowski, T.J.: Imaging Brain Dynamics Using Independent Component Analysis. *Proceedings of the IEEE* **89** (2001) 1107–1122
5. Vigário, R., Särelä, J., Jousmäki, V., Hämäläinen, M., Oja, E.: Independent Component Approach to the Analysis of EEG and MEG Recordings. *IEEE Transactions on Neural Networks* **47** (2000) 589–593
6. Tang, A., Pearlmutter, B., Malaszenko, N., Phung, D., Reeb, B.: Independent Components of Magnetoencephalography: Localization. *Neural Computation* **14** (2002) 1827–1858
7. Calhoun, V.D., Adali, T., Hansen, L.K., Larsen, J., Pekar, J.J.: ICA of functional MRI Data: An Overview. In: *Proc. 4th Int. Symp. on Independent Component Analysis and Blind Signal Separation (ICA2003)*, Nara, Japan (2003) 281–288
8. Haacke, E.M., Brown, R.W., Thompson, M.R., Venkatesan, R.: *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. 1st edn. Wiley-Interscience, New York, NY (1999)
9. Worsley, K.J., Friston, K.J.: Analysis of fMRI Time-Series Revisited — Again. *NeuroImage* **2** (1995) 173–235
10. Meinecke, F., Ziehe, A., Kawanabe, M., Müller, K.R.: A Resampling Approach to Estimate the Stability of One-Dimensional or Multidimensional Independent Components. *IEEE Transactions on Biomedical Engineering* **49** (2002) 1514–1525
11. McKeown, M.J., Hansen, L.K., Sejnowski, T.J.: Independent Component Analysis of functional MRI: What Is Signal and What Is Noise? *Current Opinion in Neurobiology* **13** (2003) 620–629
12. Himberg, J., Hyvärinen, A., Esposito, F.: Validating the independent components of neuroimaging time series via clustering and visualization. *NeuroImage* **22** (2004) 1214–1222
13. McKeown, M.J., Varadarajan, V., Huettel, S., McCarthy, G.: Deterministic and stochastic features of fMRI data: implications for analysis of event-related experiments. *Journal of Neuroscience Methods* **118** (2002) 103–113
14. Ylipaavalniemi, J., Vigário, R.: Analysis of Auditory fMRI Recordings via ICA: A Study on Consistency. In: *Proceedings of the 2004 International Joint Conference on Neural Networks (IJCNN 2004)*. Volume 1., Budapest, Hungary (2004) 249–254
15. Smith, A.M., Lewis, B.K., Ruttimann, U.E., Ye, F.Q., Sinnwell, T.M., Yang, Y., Duyn, J.H., Frank, J.A.: Investigation of Low Frequency Drift in fMRI Signal. *NeuroImage* **9** (1999) 526–533
16. Lerski, R.A., McRobbie, D.W., Straughan, K., Walker, P.M., de Certaines, J.D., Bernard, A.M.: Multi-center trial with protocols and prototype test objects for the assessment of MRI equipment. *Magnetic Resonance Imaging* **6** (1988) 201–214
17. Price, R.R., Axel, L., Morgan, T., Newman, R., Perman, W., Schneiders, N., Selikson, M., Wood, M., Thomas, S.R.: Quality assurance methods and phantoms for magnetic resonance imaging: Report of AAPM nuclear magnetic resonance Task Group No. 1. *Medical Physics* **17** (1990) 287–289
18. FastICA: MATLAB Package. In: <http://www.cis.hut.fi/research/ica/fastica>. (1998)
19. McKeown, M.J., Sejnowski, T.J.: Independent Component Analysis of fMRI Data: Examining the Assumptions. *Human Brain Mapping* **6** (1998) 368–372
20. Särelä, J., Vigário, R.: Overlearning in Marginal Distribution-Based ICA: Analysis and Solutions. *Journal of Machine Learning Research* **4** (2003) 1447–1469