Spatial independent component analysis of functional magnetic resonance imaging timeseries: characterization of the cortical components.

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Spatial independent component analysis of functional magnetic resonance imaging time-series: characterization of the cortical components

E. Formisano[a,*], F. Esposito[b], N. Kriegeskorte[a], G. Tedeschi[b], F. Di Salle[c], R. Goebel[a]

[a] *Department of Neurocognition, Faculty of Psychology, Maastricht University, Postbus 616, 6200 MD Maastricht, Netherlands*
[b] *Institute of Neurological Sciences, Second University of Naples, Naples, Italy*
[c] *Department of Neuroradiology, University of Naples "Federico II", Naples, Italy*

Abstract

Spatial independent component analysis (sICA) can be applied to human brain functional magnetic resonance imaging (fMRI) data. Here, we address the problem of identifying the "meaningful" subset in the large set of components (ICs). While this problem ultimately requires interpretation, we propose *kurtosis* of the component histogram, *spatial clustering* of the component's layout in the brain and *one-lag autocorrelation* of the time course as criteria useful in selecting components for more in-depth examination. Using our method of *cortex-based* sICA, we illustrate this selection approach by applying it to two fMRI data sets already well understood by us. The criteria in combination allow the selection of the task-related fMRI-ICs, independent of a priori information pertaining to the particular temporal structure of the experiment.
© 2002 Elsevier Science B.V. All rights reserved.

* Corresponding author. Tel.: +31-43-3884040; fax: +31-43-3884125.
  E-mail address: e.formisano@psychology.unimaas.nl (E. Formisano).

1. Introduction

In the last few years, blood-oxygenation-level-dependent (BOLD) functional magnetic resonance imaging (fMRI) has rapidly assumed a leading role among the techniques used to localize brain activity. Maps of activated brain areas are formed by detecting the small stimulus-related signal changes (1–5%) deriving from the effect of neural activity on local blood volume, flow and oxygen saturation, the so-called hemodynamic response. Typically, fMRI data are analyzed by *hypothesis-driven*, *voxel-based* statistical methods, e.g. with a general linear model (GLM). In this type of approach, a model of the hemodynamic response that is spatially invariant (i.e. equal at each voxel) is specified a priori by the experimenter and its goodness-of-fit is tested independently at each voxel [1,7]. In contrast to conventional statistical methods applied in fMRI analysis, independent component analysis (ICA) is *data-driven* and *multivariate.* The decomposition is determined solely by the intrinsic spatio–temporal structure of the data set, i.e. no a priori assumption about the time profile of the effects of interest is required.

Recently, ICA has been applied to fMRI data analysis [15,16]. The ICA-variant used is *spatial* in that the observed 4D fMRI-signals are modelled as linear "mixtures" of unknown spatially independent processes (e.g. BOLD signal changes related to the cognitive task, physiological pulsations, head movements, artifacts, etc.), each contributing to the data set with an unknown time profile. An adaptive ICA algorithm [3] is used to decompose the time series into spatial components (ICs), each having a unique time course (TC). The decomposition process maximizes the spatial statistical independence of the components, the idea being that the new representation of the data (ICs/TCs) reflects the "unmixed" configuration of the original spatial processes. In sICA, the number of spatial components serving as a linear model the 4D original data set is limited by the number of time samples acquired. As acquisition of a single fMRI volume of the brain typically takes 1–2 s and the measurement lasts for a few minutes, the total number of components that the model comprises is usually >100 and <500.

In the sICA, as proposed in [15], the entire matrix of the fMRI time series is blindly decomposed. This matrix includes signals not only from the cerebral cortex, but also from other parts of the brain, including subcortical structures, the white matter and the ventricles. The resulting decomposition, thus, also models the dynamics of the signal in these other structures. As the maximal number of spatial components is fixed (it equals the number of time samples, i.e. functional scans), the model's power of spatial (i.e. anatomical) discrimination can be enhanced by restricting the analysis to a subregion of the matrix. Whenever the cortex is the target of investigation, it is therefore advisable to restrict the sICA to the portion of the matrix that represents the cortex. Reconstructions of the cortical sheet, obtained by automatic segmentation of high spatial resolution anatomical magnetic resonance (MR) images [8,13], can be used to select the subset of the cortical time courses. The increase in spatial discriminatory power is enormous because only about 20% of the voxels of a whole brain functional data set lie within the cortex. Besides enhancing the capability of sICA to detect cortical sources (cortex-based ICA or cbICA) [6], this approach also reduces the computational load and speeds up the convergence of ICA algorithms.

As the number of ICs is often very large in spatial ICA (up to the number of time samples), a problem the user has to deal with is the selection of a "meaningful" subset from the large set of components. While this problem ultimately requires interpretation—grounded in the knowledge of an expert, certain computable properties of the obtained components and their time courses can be utilized to guide the expert's exploration of the results of the ICA.

As it is usually uninformative to order the ICs according to the contribution to the variance of the original data [16], currently adopted solutions revert to a priori information: they select the components on the basis of their time courses' correlation with the stimulation protocol [14]. Since the hallmark of ICA is *blind* decomposition, allowing the analysis of signals difficult to model a priori, a method of characterizing the fMRI-ICs independent of the stimulation protocol is desirable.

The present work aims to investigate whether three descriptive measures computed for each component are useful in selecting "meaningful" ICs in the context of fMRI of the human cerebral cortex. The three measures depend on the component's distribution of the values, on its spatial structure and on the properties of its associated TC, respectively, but not on a priori information. They are: (1) the *kurtosis* of the component's distribution of voxel values, (2) the *degree of spatial clustering* of its suprathreshold voxels, and (3) the *one-lag serial autocorrelation* of its TC. In order to test the heuristic value of these measures, we have determined them for the components obtained by cbICA of data sets well suited to conventional analysis and already well understood by us [9]. If these heuristic criteria turn out to work well in these well understood cases, they might also be helpful in selecting the meaningful components in cases where conventional analysis is not an option (e.g. because the temporal structure of the data is not known a priori).

2. fMRI experiments

fMRI data sets consisted of functional time series from a normal subject from the control group in a study of blindsight patients [9] and from three normal subjects participating in a study with a simple motor activation paradigm. Functional imaging was performed at 1.5 T (Siemens magnetom vision) using a standard head coil and a gradient-echo EPI sequence (TE $= 69$ ms, FA $= 90°$, TR $= 3000$ ms [visual study]/5000 ms [motor study], FOV $= 200 \times 200$ mm^2, voxel size $= 1.6 \times 1.6 \times 3$ mm^3, 126 scans/time series). In the first study, colored images of natural objects (fruit and vegetables) subtending $5.2 \times 5.2°$ were presented in the upper visual field $7°$ off-axis in one hemifield at a time. Each stimulation block lasted for 30 s and was repeated four times within each recording session. Fixation blocks of equal length separated stimulation blocks. Within a stimulation block, an image was shown for 1 s and was then replaced by the next image without an interstimulus interval (for details see [9]). In the second study, 10 stimulation blocks (25 s), during which subjects performed a self-paced finger-tapping task, were alternated with 10 control blocks of equal length (100 scans/time series). For all the subjects, a high-resolution T1-weighted 3D data set covering the entire brain was collected with a 3D MPRAGE

sequence (magnetization prepared rapid acquisition gradient echo, $TR = 9.7$ ms, $TE = 4$ ms, $FA = 12°$, matrix $= 256 \times 256$, thickness $= 1$ mm, number of partitions $= 170$–180, voxel size $= 1 \times 1 \times 1$ mm^3) during the same session as the functional measurement.

3. cbICA of fMRI time series[1]

cbICA of the fMRI data sets consisted of the following steps:

(1) *Reconstruction of the cortical surface.* The high-resolution T1-weighted 3D volumes were used for surface reconstruction of both cortical hemispheres of the subject. The white/gray matter border was segmented with a region-growing method preceded by inhomogeneity correction of signal intensity across space. The border of the resulting segmented subvolumes was tessellated to produce a polygon mesh representation of each cortical hemisphere [13]. The tessellation of the white/gray matter boundary of a single hemisphere yields a polygon mesh typically consisting of approximately 250,000 triangles. An iterative 3D morphing algorithm was used to move all surface vertices outwards along with their normals, such that the surface came to represent the spatial structure of the cortical gray matter. Through visual inspection, this process was halted when the surface reached the middle of the gray matter corresponding approximately to layer 4 of the cortex [8]. The resulting surface was used as the spatial reference in projecting functional data onto inflated representations of the cortex (Fig. 1a).

(2) *Selection of the cortical time courses.* The complete functional time series were interpolated to $3 \times 3 \times 3$ mm^3 resolution and co-registered to the 3D anatomical volume following procedures described elsewhere [4]. This yielded a 4D data representation (volume time course or VTC) for each functional data set that, being co-registered to the 3D anatomy, was also co-registered to the polygon mesh representing the cortical sheet. This mesh was projected into the functional data set and the voxels that were within a specified range were tagged as "cortex" (Fig. 1b). The corresponding functional time courses formed the reduced matrix of the cortical time courses [6,10].

(3) *Spatial ICA decomposition.* Let $\mathbf{X}_c$ be the $T \times M_c$ ($T =$ number of scans, $M_c =$ number of cortical time-courses) matrix of the observed cortical time courses (as defined in step 2), $\mathbf{C}$ the $N \times M_c$ matrix whose rows C_i ($i = 1, \ldots, N$) contain the spatial processes ($N \leqslant T =$ number of processes) and $\mathbf{A}$ the $T \times N$ mixing matrix whose columns A_j ($j = 1, \ldots, N$) contain the time courses of the N processes. The problem of the ICA-decomposition of fMRI time series can be formulated as the

---

[1] The proposed methods are available as a part of the BrainVoyager 2000 fMRI analysis package (www.brainvoyager.com).
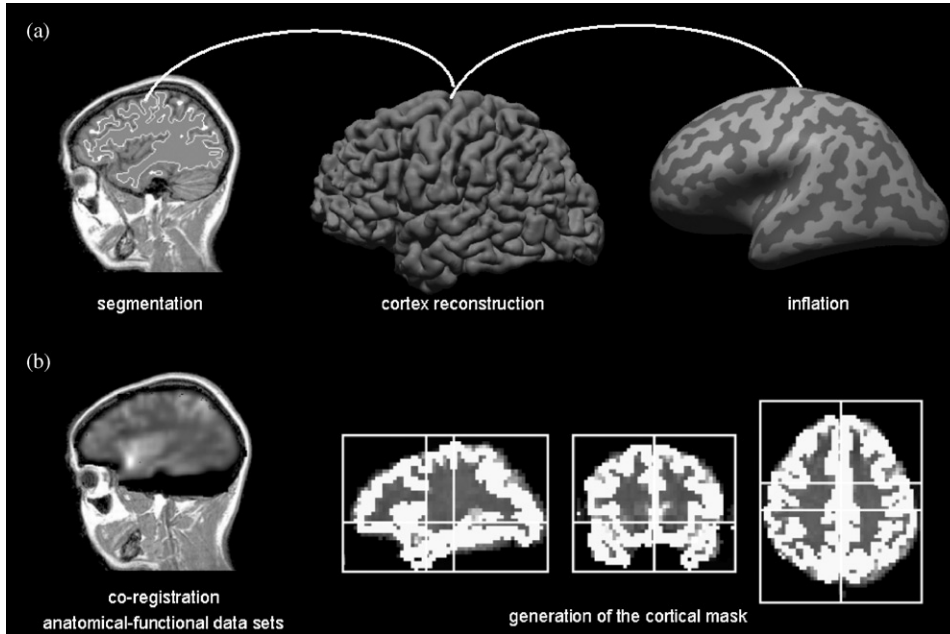
Fig. 1. The use of individual subject cortex reconstruction for visualization on an inflated representation (a) and for tagging of *cortical* voxels in the 4D matrix obtained by functional magnetic resonance imaging (b).

estimation of both matrices of the right side of the following equation:

$$\mathbf{X}_c = \mathbf{A} \cdot \mathbf{C} \tag{1}$$

under the constraint that the processes C_i are spatially independent. No a priori assumption is made about the mixing matrix A, i.e. about the time courses of the processes.

Different ICA algorithms solve this problem using different strategies (see [11] for a review). Here, $\mathbf{X}_c$ was blindly decomposed with a fixed-point ICA algorithm (FastICA, [12]) as implemented in [17]. The FastICA algorithm minimizes the mutual information of the components using a robust approximation of the negentropy as a contrast function and a fast, iterative (non-adaptive) algorithm for its maximization (for a detailed description of the FastICA algorithm, see [12]). After sphering the matrix $\mathbf{X}_c$, the hierarchical (deflation) mode of the FastICA algorithm was used and all the components ($N_1 = 126$ in the visual study and $N_2 = 100$ in the motor study) were estimated one-by-one.

Note that since these components consist of voxel values at locations linked to the cortical surface, they could directly be visualized on a folded or a morphed representation of the cortex (*cortical surface components* or CSC).

4. Characterization of the cortical surface components

After cbICA decomposition, each of the obtained CSCs was assigned values and ranks as follows:

(1) The *kurtosis* of the distribution of voxel values associated with each CSC was computed. The CSCs were ranked according to descending values of the *kurtosis*. This ranking is related to the hierarchical order of extraction of the deflation approach used in this paper for the ICA-decomposition. The use of this measure was suggested by the empirical observation that, in several runs of the ICA algorithm on functional time series collected with the same stimulation paradigm, CSCs had a super-Gaussian distribution of voxel values with values of kurtosis corresponding approximately to the same interval of the ranking [5].

(2) The spatial maps were *z*-normalized as in [16]. The number of voxels (N_{tot}) exceeding a threshold value ($|z| > 3.5$) and the size of the subset of these voxels (N_{clu}) belonging to a 3D cluster of minimum extension (100 mm^3) were computed. The CSCs were ranked according to descending values of $CLU = N_{clu}/N_{tot}$. This method aims to quantify the degree of *clustering* of the CSCs and exploits the fact that meaningful processes tend to have a well-defined spatial structure.

(3) The *one-lag serial autocorrelation coefficient* (1L) was computed for the CSCs' time courses. The CSCs were ranked according to the descending values of 1L. This method relies on the fact that potentially interesting components present a time course that is temporally structured without necessarily being correlated to the stimulation protocol [2].

Additionally, for each CSC, the root-mean-square (RMS) contribution to the original data was computed. Frequency domain analysis and multiple regression with a linear model of the hemodynamic response to the stimulation conditions was performed for the ICs' time courses.

5. Results

Fig. 2 shows the components (referred to as CSC1 and CSC2) corresponding to the cortical response to the visual stimulation (objects presented in the left and right hemifield alternately) as identified by using the a priori information of the stimulation protocol: CSC1 and CSC2 are the two components whose TCs were most highly correlated ($R = 0.76$ and 0.73) with hemodynamic predictor functions (Fig. 2b, yellow) computed on the basis of the stimulation protocol by a linear model [9]. The components were normalized and voxels with $|z|$-values > 3.5 and belonging to a 3D-cluster $> 100 \text{ mm}^3$ were color-coded. The resulting maps were projected onto the inflated representation of the subject's cortex. The brain areas integrated in these components included all early visual areas in the occipital cortex and in the ventral stream that responded to the presentation of the objects in the contralateral hemifield. Early areas of the visual processing stream were represented in a single component, while more ventral areas received the contribution of both the components (see Fig. 2a, blue patch
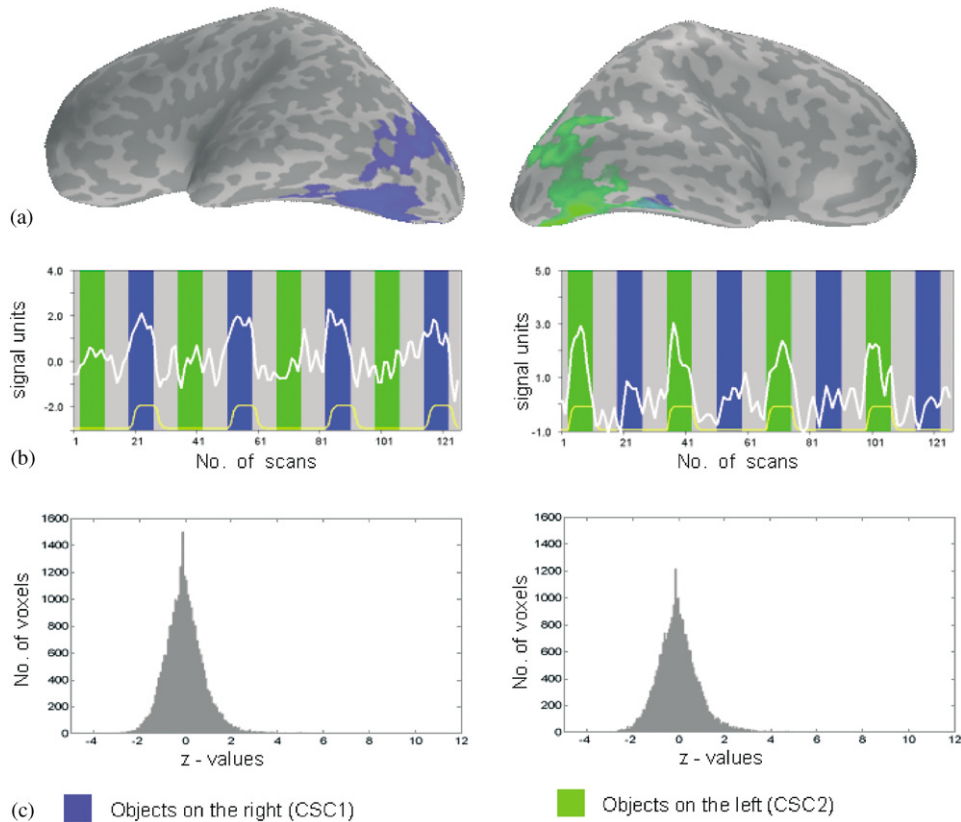
Fig. 2. The two cortical surface components whose time courses (b, white) were most highly correlated with the hemodynamic response (b, yellow) as predicted on the basis of the stimulation protocol (b, background) by a linear model. The component voxels whose values were in the tails of the distribution (c) have been color-coded on the inflated representation (a) note that, as expected from visual neuroanatomy, both components are almost completely lateralized to the hemisphere contralateral of the visual hemifield stimulated during the phases when their time courses are high.

in right hemisphere), indicating that their response was less selective to the hemifield of stimulation. The spatial locations of these areas corresponded very well to the locations of the areas identified by the multiple regression analysis (see [9] for a comparison). The power spectrum of both TCs presented prominent peaks at the frequency of the stimulus presentation, following harmonics. These results indicate that cbICA can replicate "blindly" the results of the multiple regression analysis and detect both the spatial patterns and the time courses of the task-related cortical activations.

   In Fig. 3, the correlation between all component TCs and the two predictors used in the GLM analysis (a), the RMS contribution to the original data (b) and the *kurtosis* of the 3D spatial map (c) are shown for all CSCs. In all the graphs, the CSCs are ranked according to the order in which they were obtained with the FastICA algorithm (deflation mode). By the criterion of RMS contribution, CSC1 (black in Figs. 3 and 4)
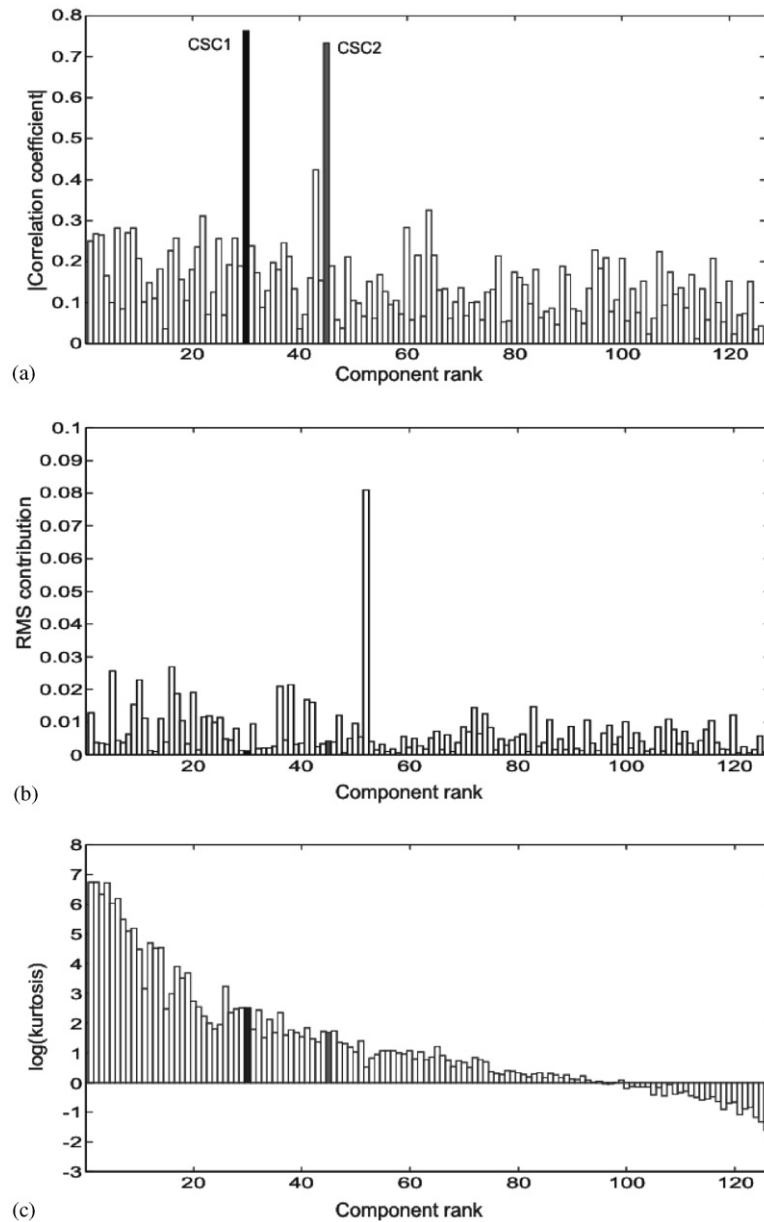
Fig. 3. The cortical surface components (CSC1 and CSC2) identified by correlation with the stimulation protocol (a) are inconspicuous by the criteria of RMS contribution (b) and kurtosis (c).

and CSC2 (gray in Figs. 3 and 4) were ranked 116th and 91st, respectively, indicating that spatial processes other than stimulus-related activation contributed most to the fMRI data sets. The CSC with the highest RMS contribution was a spatially distributed component, accounting for global changes in the signal during the scanning session. By the criterion of *kurtosis*, the selected components were ranked 23th and 41st, respectively. In the analysis of other data sets (of this and other fMRI experiments with the same number of scans and similar stimulation protocol), the *kurtosis* usually ranked the task-related CSCs in the same interval (approximately between the 21st and the 60th component). In many cases, the highly "super-Gaussian" CSCs (i.e. the components with high values of kurtosis, approximately those with ranks 1–20) represented spatially localized processes, which could usually be identified as vessels or motion-related artifacts. CSCs with low values of kurtosis presented a noise-like spatial distribution of voxel values without focal regions of activity.

In Fig. 4, the degree of clustering (CLU) of the 3D spatial maps (a) and the *one-lag serial autocorrelation* ($1L$) (b) are shown for all the CSCs. Most of the CSCs with a very high *kurtosis* also had a high value of CLU (> 0.5). Visual inspection of these CSCs revealed that one high focal cluster of voxels with very high $|z|$-values was present. High CLUs were also obtained for CSC1 and CSC2 (0.85 and 0.76, respectively). However, in this case high $|z|$-values were distributed over several clusters in the visual areas (Fig. 2). As expected, the TCs of the task-related CSCs showed high values of $1L$ (0.82 and 0.72). The highest $1L$ was obtained for one CSC with a diffuse spatial distribution. The TC of this component presented a pronounced low-frequency drift, a common artifact in fMRI data sets.

Fig. 4c shows a plot in which $1L$ values are plotted versus CLU values. In this graph, CSC1 and CSC2 are represented near the right upper corner. We would expect other potentially interesting CSCs (i.e. the CSCs with a structured TC and with several clusters of activated voxels) to be found in the same quadrant of the graph. Thus, after excluding the CSCs with very high kurtosis (1st–20th, indicated with crosses),[2] the two nearest CSCs to the task-related components (indicated by circles) were inspected. Each of them formed a symmetric pattern over the two hemispheres and their TCs suggested transient activation. While the first represented a bilateral region of the occipital cortex, the second included clusters of voxels in the anterior cingulum and in the frontal cortex. Interestingly, the TC of the first of the two indicated occipital activations during the periods when the subject saw only the fixation cross. Although the exploratory nature of this type of analysis does not allow for quantitative inferences, these CSCs might reflect weakly task-related or -unrelated cognitive activity of the subject during the scanning session.

In Fig. 5, the same $1L$–CLU plot as in Fig. 4c is shown for the functional time series collected during the motor task. For each of the three subjects, the suggested method was capable of identifying task-related components, independent of the stimulation protocol. In all the cases, the CSCs represented near the right upper corner of the graph included primary motor, pre-motor and supplementary motor areas, while the

---

[2]Note that this exclusion was carried out on the basis of a visual inspection of the components. It is possible that, in other experiments, potentially meaningful components are ranked in this same interval.
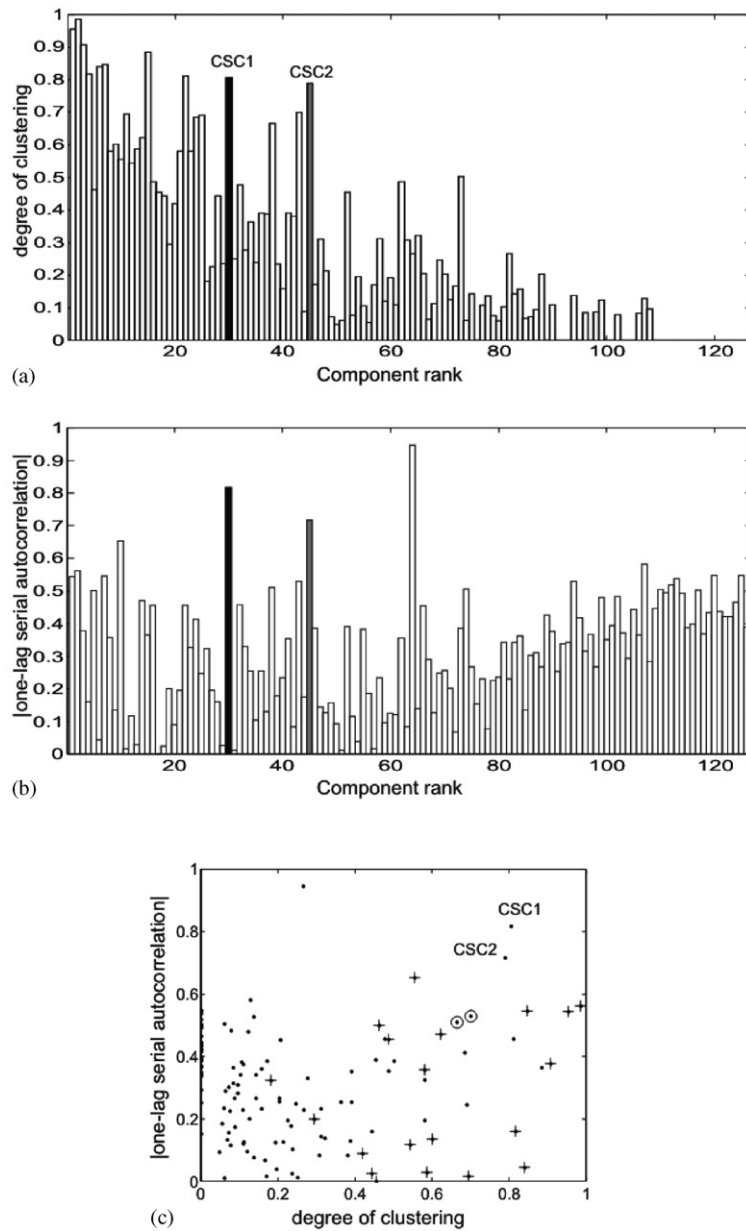
Fig. 4. The combination (c) of degree of clustering (a) as a measure of spatial structure and one-lag serial autocorrelation (b) as a measure of temporal structure is well-suited to detect the "meaningful" components. In (c), crosses (+) mark high-kurtosis components (ranks 1–20), circles (o) mark potentially "meaningful" components that merit more in-depth examination.
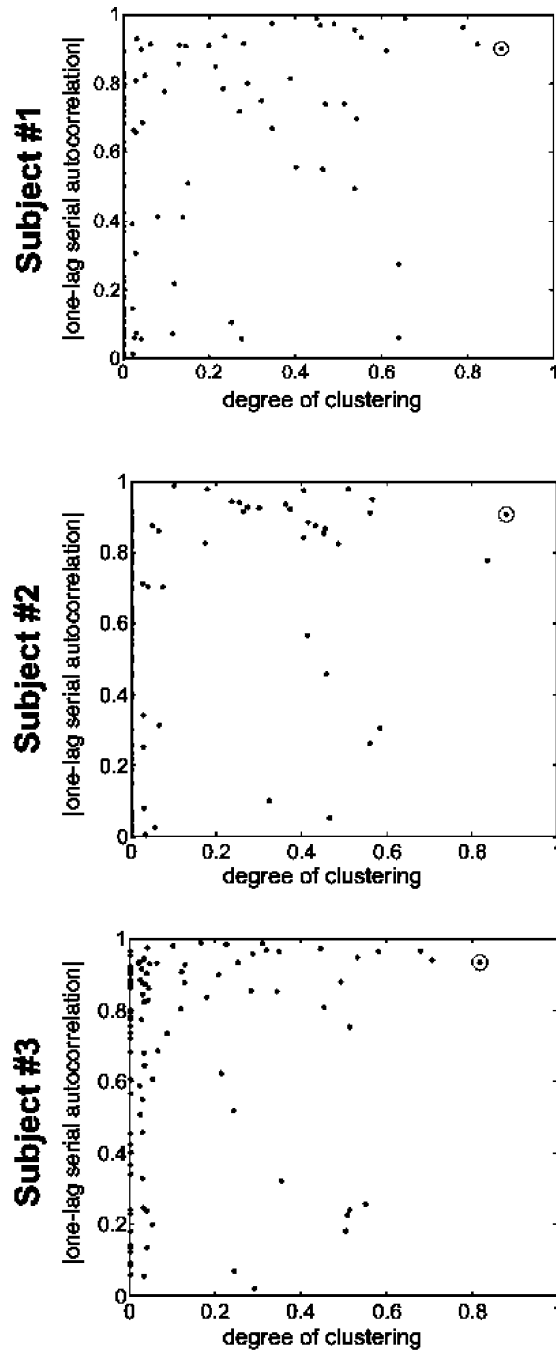
Fig. 5. For each of the three subjects of the motor experiment, the combination of degree of clustering as a measure of spatial structure and one-lag serial autocorrelation as a measure of temporal structure identifies the task-related component (marked by a circle) independent of the stimulation protocol.

associated time courses were highly correlated ($R = 0.71$, 0.76, and 0.69, respectively) with hemodynamic predictor functions.

6. Conclusions

In the present paper we have characterized fMRI-ICs, as obtained by cortex-based ICA, on the basis of several parameters that depend on the spatial structure of the 3D voxel values' distribution or on the intrinsic properties of the associated TCs. We have shown that a suitable combination of these parameters allows the selection of potentially "meaningful" fMRI-ICs, independent of the correlation of component-TCs with the stimulation protocol. Although validation of this approach using a larger number of fMRI data sets is needed, we believe that this approach is very useful in the analysis of fMRI experiments in which the hemodynamic response is difficult to model (e.g. event-related designs during complex cognitive tasks) or in which the events of interest are unpredictable (e.g. hallucinations, epileptic seizures).

References

[1] P.A. Bandettini, A. Jesmanowicz, E.C. Wong, J.S. Hyde, Processing strategies for time-course data sets in functional MRI of the human brain, Magn. Reson. Med. 30 (1993) 161–173.

[2] R. Baumgartner, R. Somorjai, R. Summers, W. Richter, L. Ryner, Novelty indices: identifiers of potentially interesting time-courses in functional MRI data, Magn. Reson. Imaging 18 (2000) 845–850.

[3] A.J. Bell, T.J. Sejnowski, An information-maximization approach to blind separation and blind deconvolution, Neural Comput. 7 (1995) 1129–1159.

[4] T. Dierks, D.E.J. Linden, M. Jandl, E. Formisano, R. Goebel, H. Lanfermann, W. Singer, Activation of Heschl's gyrus during auditory hallucinations, Neuron 22 (1999) 615–621.

[5] F. Esposito, E. Formisano, S. Cirillo, R. Elefante, G. Tedeschi, R. Goebel, F. Di Salle, Critera for rank ordering of fMRI independent components, NeuroImage 13 (2001) S114.

[6] E. Formisano, F. Esposito, F. Di Salle, R. Goebel, Cortex-based independent component analysis of fMRI time-series, NeuroImage 13 (2001) S119.

[7] K.J. Friston, A.P. Holmes, K.J. Worsley, J.B. Poline, C.D. Frith, R.S.J. Frackowiak, Statistical parametric maps in functional imaging: a general linear approach, Human Brain Map 2 (1995) 189–210.

[8] R. Goebel, D. Khorram-Sefat, L. Muckli, H. Hacker, W. Singer, The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery, Eur. J. Neurosci. 10 (1998) 1563–1573.

[9] R. Goebel, L. Muckli, F.E. Zanella, W. Singer, P. Stoerig, Sustained extrastriate cortical activation without visual awareness revealed by fMRI studies of hemianopic patients, Vision Res. 41 (2001) 1459–1474.

[10] R. Goebel, W. Singer, Cortical surface-based statistical analysis of functional magnetic resonance imaging data, NeuroImage 9 (1999) S64.

[11] A. Hyvärinen, A survey on independent component analysis, Neural Comput. Surv. (1999) 94–128.

[12] A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, IEEE Trans. Neural Networks 10 (1999) 626–634.

[13] N. Kriegeskorte, R. Goebel, An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical MR volumes, NeuroImage 14 (2001) 329–346.

[14] M.J. McKeown, Detection of consistently task-related activations in fMRI data with hybrid independent component analysis, NeuroImage 11 (2000) 24–35.

[15] M.J. Mc Keown, T.P. Jung, S. Makeig, G.G. Brown, S.S. Kindermann, T.W. Lee, T.J. Sejnowski, Spatially independent activity patterns in functional MRI data during the Stroop color-naming task, Proc. Natl. Acad. Sci. USA 95 (1998) 803–810.

[16] M.J. Mc Keown, S. Makeig, G.G. Brown, T.P. Jung, S.S. Kindermann, A.J. Bell, T.J. Sejnowski, Analysis of fMRI data by blind separation into independent spatial components, Human Brain Map 6 (1998) 160–168.

[17] The FastICA MATLAB package (http://www.cis.hut.fi/projects/ica/fastica/).

Elia Formisano is Ph.D. in Biomedical Engineering and, at present, Assistant Professor in the University of Maastricht (Dept. of Cognitive Neuroscience/Faculty of Psychology). His main research interests are the development of methods for the analysis and representation of functional magnetic resonance imaging data and the study of the functional organization of the human auditory cortex.

Fabrizio Esposito is a Ph.D. student in the Department of Neurological Sciences at the Second University of Naples. His main research interests are the methods for the non-inferential analysis and representation of functional magnetic resonance imaging data.

Nikolaus Kriegeskorte is a Ph.D. student in the Department of Cognitive Neuroscience at Maastricht University. He studied psychology and computer science at the University of Cologne. His interests include psychophysics, fMRI, and network modeling of high-level functions of the primate visual system as well as fMRI methodology.

Gioacchino Tedeschi is full Professor of Neurology in the Department of Neurological Sciences of the Second University of Naples. His main research interests are the Magnetic Resonance methods for the spectroscopic and functional analysis of the brain.

Francesco Di Salle is a MD and, at present, Assistant Professor at the University "Federico II" of Naples (Dept. of Neurological Sciences/Faculty of Medicine). His main research interests are the methods for the analysis and representation of functional magnetic resonance imaging data and the fMRI of the auditory cortex.

Rainer Goebel is a full professor for Cognitive Neuroscience in the Faculty of Psychology of Maastricht University and a research fellow in the F.C. Donders Centre for Cognitive Neuroimaging in Nijmegen, Netherlands. He is currently combining functional MRI and EEG/MEG in order to noninvasively measure both the topographic distribution and the precise timing of neuronal processes in humans.