Model–Free Functional MRI Analysis Using Improved Fuzzy Cluster Analysis Techniques

Oliver Lange[1,2], Anke Meyer–Bäse[1], Axel Wismüller[1,2],

[1]Department of Electrical and Computer Engineering,
Florida State University, Tallahassee, Florida 32310-6046 USA

[2]Department of Clinical Radiology, Ludwig–Maximilians University,
Munich 80336, Germany

Monica Hurdal, DeWitt Sumners, and

Department of Mathematics,
Florida State University, Tallahassee, Florida 32306-4510 USA

Dorothee Auer

Max Planck Institute of Psychiatry, Munich 80804, Germany

ABSTRACT

Conventional model–based or statistical analysis methods for functional MRI (fMRI) are easy to implement, and are effective in analyzing data with simple paradigms. However, they are not applicable in situations in which patterns of neural response are complicated and when fMRI response is unknown. In this paper the Gath-Geva algorithm is adapted and rigorously studied for analyzing fMRI data. The algorithm supports spatial connectivity aiding in the identification of activation sites in functional brain imaging. A comparison of this new method with the fuzzy n–means algorithm, Kohonen's self-organizing map, fuzzy n–means algorithm with unsupervised initialization, minimal free energy vector quantizer and the "neural gas" network is done in a systematic fMRI study showing comparative quantitative evaluations. The most important findings in this paper are: (1) the Gath–Geva algorithms outperforms for a large number of codebook vectors all other clustering methods in terms of detecting small activation areas, and (2) for a smaller number of codebook vectors the fuzzy n–means with unsupervised initialization outperforms all other techniques. The applicability of the new algorithm is demonstrated on experimental data.

Keywords: Gath-Geva algorithm, fuzzy n–means algorithm, minimal free energy vector quantization, "neural gas" network, self–organizing map, fMRI

1. INTRODUCTION

Functional magnetic resonance imaging with high temporal and spatial resolution represents a powerful technique for visualizing rapid and fine activation patterns of the human brain.[1–5] As is known from both theoretical estimations and experimental results,[4,6,7] an activated signal variation appears very low on a clinical scanner. This motivates the application of analysis methods to determine the response waveforms and associated activated regions. Generally, these techniques can be divided into two groups: Model-based techniques require prior knowledge about activation patterns, whereas model-free techniques do not. However, model-based analysis methods impose some limitations on data analysis under complicated experimental conditions. Therefore, analysis methods that do not rely on any assumed model of functional response are considered more powerful and relevant. There are two kinds of model-free methods. The first method, principal component analysis (PCA)[8,9] or independent component analysis (ICA),[10,11] transforms original data into high-dimensional vector space to separate functional response and various noise sources from each other.

Intelligent Computing: Theory and Applications II, edited by Kevin L. Priddy, Proc. of
SPIE Vol. 5421 (SPIE, Bellingham, WA, 2004) · 0277-786X/04/$15 · doi: 10.1117/12.541778

The second method, fuzzy clustering analysis[12–15] or self-organizing map,[15–17] attempts to classify time signals of the brain into several patterns according to temporal similarity among these signals.

In this paper, we propose to employ the Gath–Geva algorithm to increase the analysis power without sacrifying efficiency. In a systematic manner, we will compare and evaluate the results obtained based on this new approach with the traditional Kohonen's self–organizing map (SOM), the fuzzy n–means algorithms, the "neural gas" network, the fuzzy n–means algorithm with unsupervised initialization, and with the minimal free energy vector quantization (VQ).[15]

2. THE CLUSTERING ALGORITHMS

Functional organization of the brain is based on two complementary principles, localization and connectionism. Localization means that each visual function is performed mainly by a small set of the cortex. Connectionism, on the other hand, expresses that the brain regions involved in a certain visual cortex function are widely distributed, and thus the brain activity necessary to perform a given task may be the functional integration of activity in distinct brain systems. It is important to stress out that in neurobiology the term "connectionism" is used in a different sense that that used in the neural network terminology.

Cluster analysis groups image pixels together based on the similarity of their intensity profile in time. In the clustering process, a time course with n points is represented by one point in an n–dimensional Euclidean space which is subsequently partitioned into clusters based on the proximity of the input data.

2.1. FUZZY CLUSTERING ALGORITHMS

Traditional statistical classifiers assume that the pdf for each class is known or must somehow be estimated. Another problem is posed by the fact that sometimes clusters are not compact but shell-shaped. A solution to this problem is given by fuzzy clustering algorithms, a new classification paradigm intensively studied during the past three decades. The main difference between traditional statistical classification techniques and fuzzy clustering techniques is that in the fuzzy approaches an input vector belongs *simultaneously* to more than one cluster, while in statistical approaches it belongs *exclusively* to only one cluster.

Usually, clustering techniques are based on the optimization of a cost or objective function J. This predefined measure J is a function of the input data and of an unknown parameter vector set $\mathbf{L}$. Throughout this chapter, we will assume that the number of clusters n is predefined and fixed.

A successful classification is based on estimating the parameter $\mathbf{L}$ such that the cluster structure of the input data is as good as possible determined. It is evident that this parameter depends on the cluster's geometry. Compact clusters are pretty well described by a set of n points $\mathbf{L}_i \in \mathbf{L}$ where each point describes such a cluster. Spherical clusters have two distinct parameters describing the center and the radius of the cluster. Thus the parameter vector set $\mathbf{L}$ is replaced by two new parameter vector sets, $\mathbf{V}$ describing the centers of the clusters, and $\mathbf{R}$ describing the radii of the clusters.

In the following, we will review the most important fuzzy clustering techniques, and show their relationship to nonfuzzy approaches.

2.1.1. FUZZY n–MEANS ALGORITHM

We consider here compact clusters completely described by a point representative $\mathbf{L}_i$.

Let $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ define the data set.

The following assumptions are made:

- X has a cluster substructure described by the fuzzy partition $P = \{A_1, \ldots, A_n\}$.
- n is the number of known subclusters in X.

In order to monitor the convergence of the algorithm, the $n \times p$ partition matrix $\mathbf{Q}^i$ is introduced to describe each fuzzy partition P^i at the ith iteration and is used to determine the distance between two fuzzy partitions. The matrix $\mathbf{Q}^i$ is defined as

$$\mathbf{Q}^i = \mathbf{U} \quad \text{at} \quad \text{iteration} \quad \text{i} \tag{1}$$

The termination criterion for iteration m is given by

$$d(P^m, P^{m-1}) = ||\mathbf{Q}^m - \mathbf{Q}^{m-1}|| < \epsilon \tag{2}$$

where ϵ defines the admissible error and $||\cdot||$ is any vector norm.

An algorithmic description of the fuzzy n–means algorithm is given below:

1. **Initialization:** Choose the number n of subclusters in X and the termination criterion ϵ. P^1 is selected as a random fuzzy partition of X having n atoms. Determine $\mathbf{Q}^1$, the matrix representation of P^1.

2. **Adaptation, part I:** Determine the cluster prototypes $\mathbf{L}_i, i = 1, \ldots, n$ using

$$\mathbf{L}_i = \frac{1}{\sum_{j=1}^{p} u_{ij}^2} \sum_{j=1}^{p} u_{ij}^2 \mathbf{x}_j \tag{3}$$

3. **Adaptation, part II:** Determine a new fuzzy partition P^2 of X using the following rules:

$$I_j = \varnothing \Rightarrow u_{ij} = \frac{1}{\sum_{k=1}^{n} \frac{d^2(\mathbf{x}_j, \mathbf{L}_i)}{d^2(\mathbf{x}_j, \mathbf{L}_k)}}, \quad \forall 1 \leq i \leq n; \quad 1 \leq j \leq p \tag{4}$$

and

$$I_j \neq \varnothing \Rightarrow u_{ij} = 0, \forall i \in I_j \tag{5}$$

and arbitrarily assign $\sum_{i \in I_j} u_{ij} = 1$.

Determine $\mathbf{Q}^2$, the matrix representation of the fuzzy partition P^2.

4. **Continuation:** If the difference between two successive partitions is smaller than a predefined threshold $||\mathbf{Q}^1 - \mathbf{Q}^2|| < \epsilon$, then stop. Else set $P^1 = P^2, \mathbf{Q}^1 = \mathbf{Q}^2$ and go to step 2.

2.1.2. THE GATH–GEVA ALGORITHM

A major problem arises when fuzzy clustering is performed in real–world tasks: the necessary cluster number, their locations, shapes and densities are usually not known beforehand, and thus a cluster validity criterion has to be employed to determine the optimal number of clusters.

To allow the detection of cluster shapes ranging from spherical to ellipsoidal clusters, different metrics have to be used. Usually, an adaptive metric is used. We define a distance metric $d(\mathbf{x}_j, \mathbf{L}_i)$ from the data point $\mathbf{x}_j$ to the cluster prototype $\mathbf{L}_i$ as

$$d^2(\mathbf{x}_j, \mathbf{L}_i) = (\mathbf{x}_j - \mathbf{L}_i)^T \mathbf{M}_i (\mathbf{x}_j - \mathbf{L}_i) \tag{6}$$

where $\mathbf{M}_i$ is a symmetric and positive definite shape matrix and adapts to the clusters' shape variations. For the fuzzy n–means algorithm we have $\mathbf{M} = \mathbf{I}$.

For hyperellipsoidal clusters, as well as in the presence of variable cluster densities and unequal number of data points in each cluster, an "exponential" distance measure, $d_e^2(\mathbf{x}_j, \mathbf{L}_i)$, based on maximum likelihood estimation is defined. This distance is used to determine the posterior probability $h(i|\mathbf{x}_j)$ which represents the probability of selecting the ith cluster given the jth feature vector:

$$h(i|\mathbf{x}_j) = \frac{1/d_e^2(\mathbf{x}_j, \mathbf{L}_i)}{\sum_{k=1}^{n} 1/d_e^2(\mathbf{x}_j, \mathbf{L}_k)} \tag{7}$$

and the exponential distance is given by

$$d_e^2(\mathbf{x}_j, \mathbf{L}_i) = \frac{\sqrt{|\mathbf{M}_i|}}{P_i} e^{[(\mathbf{x}_j - \mathbf{L}_i)^T \mathbf{M}_i (\mathbf{x}_j - \mathbf{L}_i)/2]} \tag{8}$$

where p_i is the a priori probability of selecting the ith cluster and is given by:

$$P_i = \frac{1}{n} \sum_{j=1}^{n} h(i|\mathbf{x}_j) \tag{9}$$

and the shape matrix $\mathbf{M}_i$ is given by

$$\mathbf{M}_i = \frac{\sum_{j=1}^{n} h(i|\mathbf{x}_j)(\mathbf{x}_j - \mathbf{L}_i)(\mathbf{x}_j - \mathbf{L}_i)^T}{\sum_{j=1}^{n} h(i|\mathbf{x}_j)} \tag{10}$$

This algorithm seeks due to this exponential distance an optimum in a narrow local region. Its major advantage is obtaining good partition results in cases of unequally variable features and densities but only when the starting cluster prototypes are properly chosen.

Thus it's wise to use a two–step classification strategy: in a first step, we have to perform an unsupervised search of cluster prototypes based on the fuzzy n–means algorithm, and in the second step the optimal fuzzy partition is being carried out with the fuzzy maximum likelihood estimation.

2.2. KOHONEN'S SELF–ORGANIZING MAP

Kohonen's self–organizing map generates nodes on a two–dimensional lattice in which the distribution of these nodes corresponds to the proximity of their associated node patterns in the signal intensity space. The benefits of this clustering technique are: (1) if started with an adequate number of neurons, it can find distinctive features in the data even if they are less predominant, and (2) the emerging node patterns are ordered according to their proximity properties in the data space. This topology–preserving technique enables the forming of superclusters by fusing nodes, and thus provides a way to visualize high–dimensional data sets. Its advantages in analyzing fMRI data were demonstrated in.[16]

2.3. MINIMAL FREE ENERGY VECTOR QUANTIZATION

Another proven tool for the analysis of fMRI time-series is given by a clustering technique, the so–called minimal free energy vector quantization.[15] This clustering procedure identifies groups of pixels sharing similar properties of signal dynamics, and thus enables the interpretation of the physiological part of the experiment. The main differences between SOM and minimal free energy vector quantization were in[15] excellent pointed out: (1) the hierarchical and multiresolution aspect of data analysis, (2) monitoring based on different control parameters (free energy, entropy) facilitates straightforward cluster splitting, and (3) the learning rule based on a stochastic gradient descent on an explicitely given error function.

2.4. "NEURAL GAS" NETWORK

The "neural-gas" algorithm[18] is an efficient approach which, applied to the task of vector quantization, (1) converges quickly to low distortion errors, (2) reaches a distortion error E lower than that from Kohonen's feature map, and (3) at the same time obeys a gradient descent on an energy surface.

Instead of using the distance $||\mathbf{x} - \mathbf{w_i}||$ or of using the arrangement of the $||\mathbf{w_i}||$ within an external lattice, it utilizes a neighborhood–ranking of the reference vectors $\mathbf{w_i}$ for the given data vector $\mathbf{x}$. The adaptation of the reference vectors is given by

$$\Delta \mathbf{w_i} = \epsilon e^{-k_i(\mathbf{w}, \mathbf{w_i}/\lambda)}(\mathbf{x} - \mathbf{w_i}) \quad i = 1, \cdots, N \tag{11}$$

The step size $\epsilon \in [0, 1]$ describes the overall extent of the modification and k_i is the number of the closest neighbors of the reference vector $\mathbf{w_i}$. λ is a characteristic decay constant.

In[18] was shown that the average change of the reference vectors corresponds to an overdamped motion of particles in a potential that is given by the negative data point density. Superimposed on the gradient of this potential is a "force", which points toward the direction of the space where the particle density is low. This "force" is the result of a repulsive coupling between the particles (reference vectors). In its form it resembles an entropic force and tends to homogeneously distribute the particles (reference vectors) over the input space, like in case of a diffusing gas. This suggests the name for the "neural-gas" algorithm. It's interesting also to mention that the reference vectors change their locations slowly but permanently and, therefore, pointers that are neighboring at an early stage of the adaptation procedure might not be neighboring anymore at a more advanced stage. Connections that have not been refreshed for a while die out and are removed.

Another important feature of the presented algorithm compared to Kohonen algorithm is that it doesn't require a prespecified graph (network). In addition, it can produce topologically preserving maps, which is only possible if the topological structure of the graph matches the topological structure of the data manifold. In cases, however, where it is not possible to a priori determine an appropriate graph, for example, in cases where the topological structure of the data manifold is not known a priori or is too complicated to be specified, Kohonen's algorithm necessarily fails in providing perfectly topology preserving maps.

Applied to fMRI, it was shown in[19] that this approach is independent of the stimulation paradigm.

3. RESULTS AND DISCUSSION

FMRI data were recorded from six subjects (3 female, 3 male, age 20-37) performing a visual task. In five subjects, five slices with 100 images (TR/TE=3000/60msec) were acquired with five periods of rest and five photic simulation periods with rest. Simulation and rest periods comprised 10 repetitions each, i.e. 30s. Resolution was $3 \times 3 \times 4$ mm. The slices were oriented parallel to the calcarine fissure. Photic stimulation was performed using an 8 Hz alternating checkerboard stimulus with a central fixation point and a dark background with a central fixation point during the control periods.[15] The first scan was discarded for remaining saturation effects. Motion artifacts were compensated by automatic image alignment (AIR,[20]).

The clustering results were evaluated by (1) task–related activation maps, (2) associated time–courses and (3) ROC curves. Cluster assignment maps represent cluster membership maps obtained based on a minimal distance criterion in the pixel time course space. For the fMRI data, a comparative quantitative evaluation among the six clustering techniques, SOM, "neural gas" (NG) network, fuzzy n–means algorithm (FVQ), Gath–Geva algorithm (GGA), fuzzy n–means with unsupervised initialization (KVQ), and minimal free energy (MFE) VQ, was performed.

3.1. ESTIMATION OF THE CLUSTERING MODEL

To decide to what extent clustering techniques of fMRI time–series depend on the employed algorithm, we have first to look at the optimal number of code vectors determined by each algorithm.

In the following we will give the set parameters. For SOM we choose: (1) an one–dimensional lattice, and (2) the maximal number of itera tions. For "neural gas" network we choose: (1) the learning parameters $\epsilon_i = 0.5$

and $\epsilon_f = 0.005$, and (2) the lattice parameters λ_i equals half the number of classes and $\lambda_f = 0.01$ and (3) the maximal number of iterations equals 1000. For the Gath–Geva algorithm and the other fuzzy algorithms, we choose the fuzzy factor=1.05, and the maximal number of iterations equals 120. And last, for minimal free energy VQ we set: (1) neurons' initialization with principal components, (2) learning parameter $\rho_{final} = 0.01$ and updating based on a linear annealing scheme, and (3) the maximal number of iterations equals 100.

Therefore, it is significant to find a fixed number of CVs that can theoretically predict new observations in same conditions, assuming the basic ICA model actually holds. To do so, we compared the six proposed algorithms for 8, 9, and 16 components in terms of ROC analysis using correlation map with a chosen threshold of 0.4 and 0.6.

The obtained results are plotted in Figures 1 and 2. The Gath–Geva algorithms outperforms for CV=16 and both chosen thresholds all other clustering methods. For CV=8,9 the fuzzy n–means with unsupervised initialization (KVQ) outperforms all other techniques.

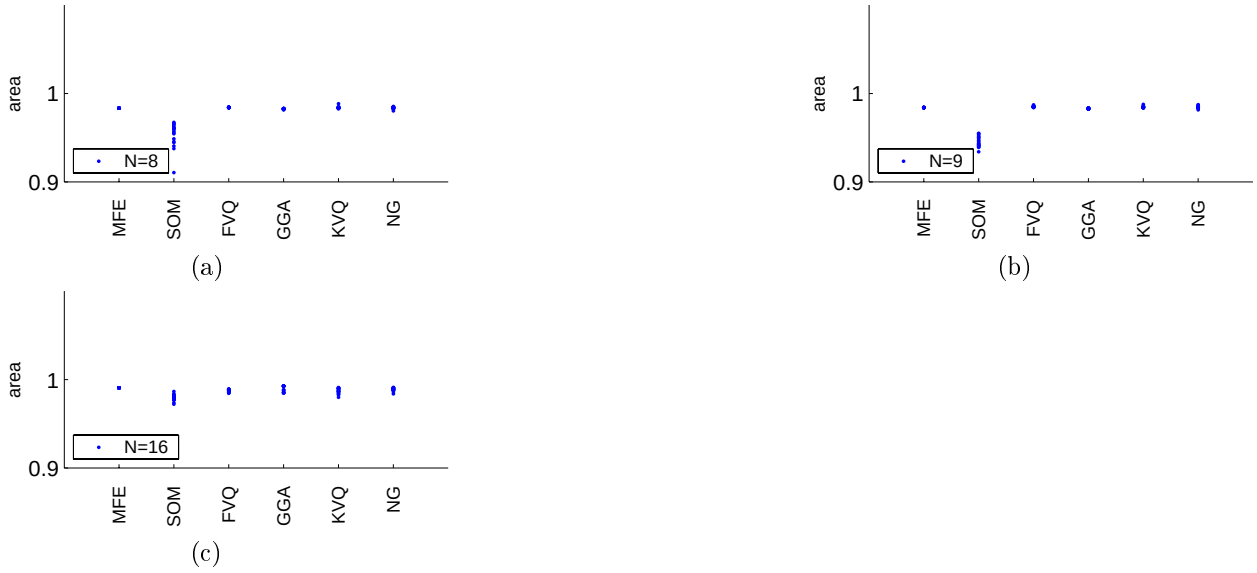

Figure 1. Results of the comparison between minimal free energy VQ (MFE), Kohonen's map (SOM), fuzzy n–means algorithm (FVQ), Gath–Geva algorithm (GGA), fuzzy n–means with unsupervised initialization (KVQ), and "neural gas" (NG) algorithm on fMRI data. Spatial accuracy of cluster analysis maps is assessed by ROC analysis using correlation map with a chosen threshold of 0.4. The number of chosen codebook vectors for all techniques is in (a): CV=8, (b): CV=9, and (c): CV=16.

The clustering results for the new method is shown in Figures 3 and 4. Figure 3 illustrates the so–called assignment maps where all the pixels belonging to a specific cluster are highlighted. The assignment between a pixel and a specific cluster is given by the minimum distance between the pixel and a CV from the established codebook. On the other hand, each CV shown in Figures 4 can be viewed as the cluster–specific weighted average of all pixel time courses.

3.2. CHARACTERIZATION OF TASK–RELATED EFFECTS

For all subjects, and runs, unique task–related activation maps and associated time–courses were obtained by all proposed techniques.

Figure 5 shows for 8, 9 and 16 CVs the component time course most closely associated with the visual task for the Gath–Geva algorithm. It becomes evident that the more CVs we take into account, the higher the correlation coefficient becomes.

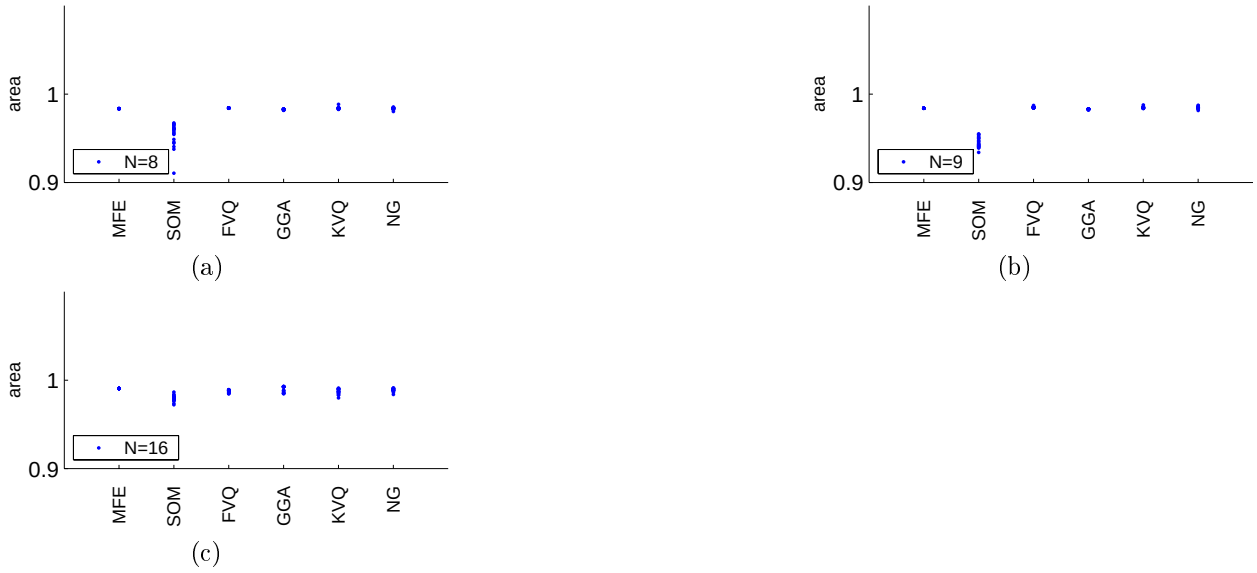

Figure 2. Results of the comparison between minimal free energy VQ (MFE), Kohonen's map (SOM), fuzzy n–means algorithm (FVQ), Gath–Geva algorithm (GGA), fuzzy n–means with unsupervised initialization (KVQ), and "neural gas" (NG) algorithm on fMRI data. Spatial accuracy of cluster analysis maps is assessed by ROC analysis using correlation map with a chosen threshold of 0.6. The number of chosen codebook vectors for all techniques is in (a): CV=8, (b): CV=9, and (c): CV=16.

4. CONCLUSION

In the present paper, we have experimentally compared proven clustering algorithms, the SOM, the minimal free energy VQ, the "neural gas" network, the fuzzy n–means algorithm, and the fuzzy n–means algorithm with unsupervised initialization with a powerful fuzzy algorithm, the Gath–Geva algorithm. The goal of the paper was to demonstrate that unsupervised clustering techniques represent an useful strategy for the analysis of time–courses from fMRI data sets. The increasing cluster resolution proved to reveal extremely well the structure of the data set. It has been shown that the Gath–Geva algorithms outperforms for a large number of codebook vectors all other clustering methods. For a smaller number of codebook vectors the fuzzy n–means with unsupervised initialization outperforms all other techniques. All unsupervised clustering techniques can be employed to determine artifacts, and thus improve the experimental environment. The applicability of the new algorithm is demonstrated on experimental data. The proposed techniques can be applied to various nonmedical problems, such as remote sensing applications enabling a better understanding for example of the Earth Data System.

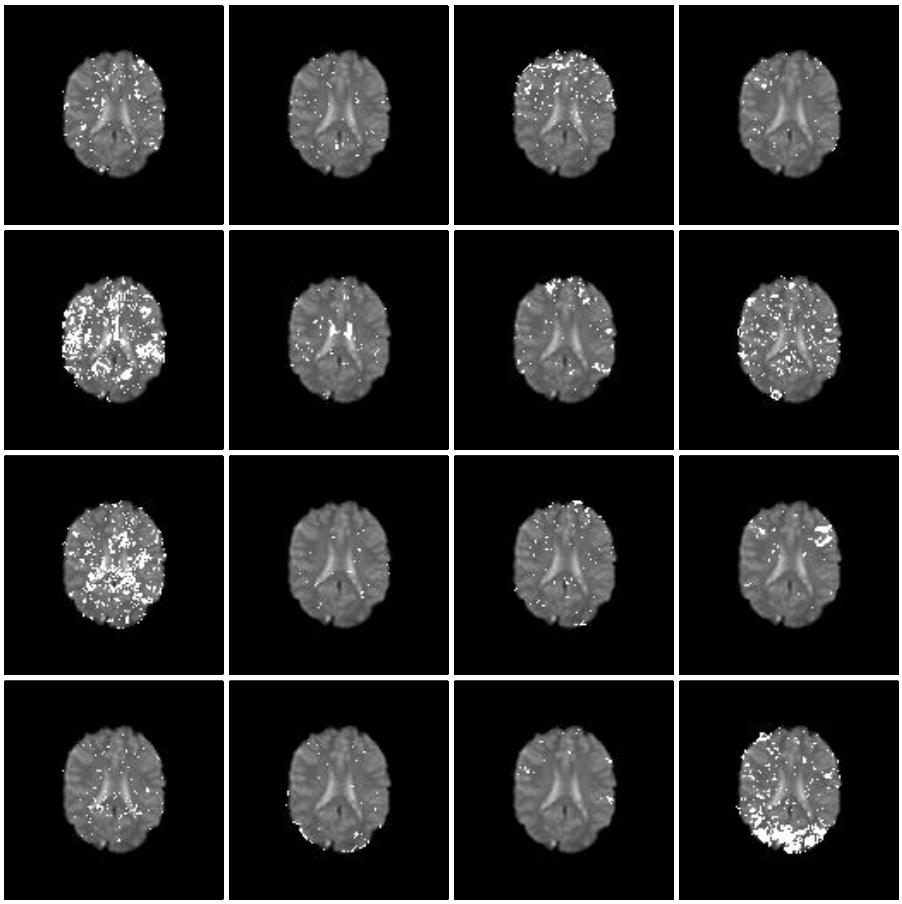

Figure 3. Cluster assignment maps for a visual stimulation fMRI experiment and obtained for 16 CVs for the Gath–Geva algorithm.

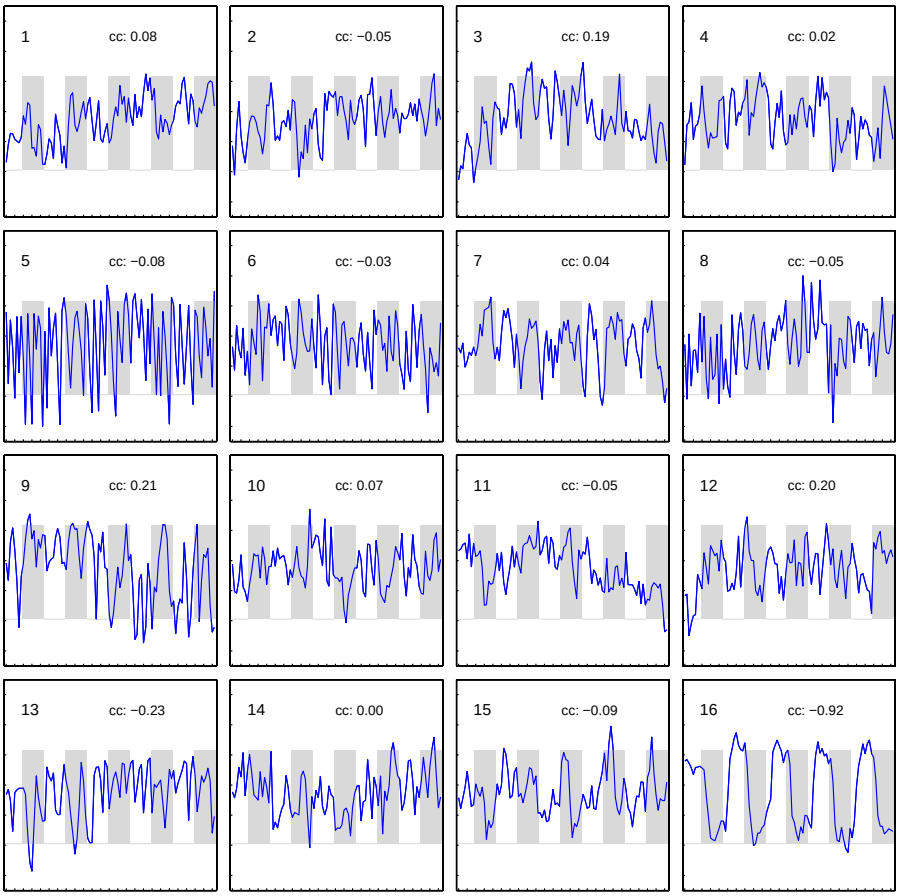

Figure 4. Associated codebook vectors for a visual stimulation fMRI experiment and obtained for 16 CVs for the Gath–Geva algorithm: (a) cluster assignment maps and (b) associated ICs.

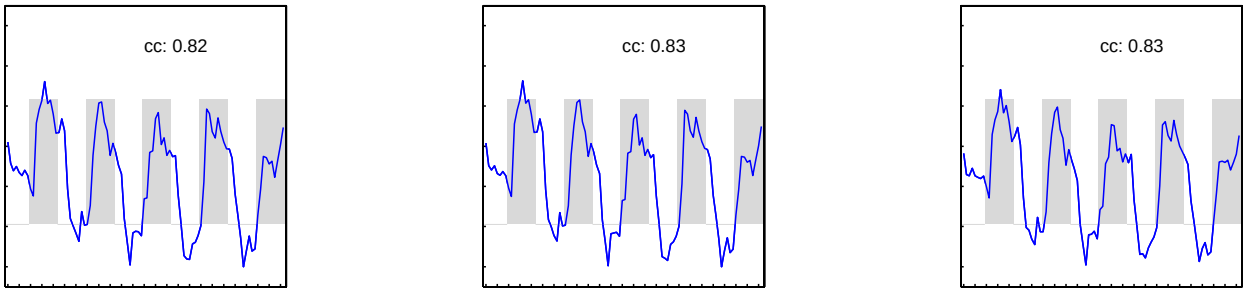

Figure 5. Computed reference functions for the Gath–Geva algorithm for CVs=8,9 and 16. The found correlation coefficients are: $r = 0.82$ for CV=8, $r = 0.83$ for CV=9, and $r = 0.83$ for CV=16.

REFERENCES

1. P. Bandettini, E. Wong, R. Hinks, R. Tikofski, and J. Hyde, "Time course epi of human brain function during task activation," *Magn. Reson. Med.* **25**, pp. 390–397, 8 1992.
2. J. Frahm, K. Merboldt, and W. Hanicke, "Functional mri of human brain activation at high spatial resolution," *Magn. Reson. Med.* **29**, pp. 139–144, 8 1992.
3. K. Kwong, "Functinal magnetic resonance imaging with echo planar imaging," *Magn. Reson. Q.* **11**, pp. 1–20, 8 1992.
4. K. Kwong, J. Belliveau, and D. Chesler, "Dynamic magnetic resonance imaging of human brain activity during primary sensor stimulation," *Proceedings of the National Academy of Science* **89**, pp. 5675–5679, 8 1992.
5. S. Ogawa, D. Tank, and R. Menon, "Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging," *Proceedings of the National Academy of Science* **89**, pp. 5951–5955, 8 1992.
6. J. Boxerman, P. Bandettini, K. Kwong, and J. Baker, "The intravsacular contribution to fmri signal change: Monte carlo modeling and diffusion-weighted studies in vivo," *Magn. Reson. Med.* **34**, pp. 4–10, 8 1995.
7. S. Ogawa, T. Lee, and B. Barrere, "The sensitivity of magnetic resonance image signals of a rat brain to changes in the cerebral venous blood oxygenation activation," *Magn. Reson. Med.* **29**, pp. 205–210, 8 1993.
8. J. Sychra, P. Bandettini, N. Bhattacharya, and Q. Lin, "Synthetic images by subspace transforms i. principal components images and related filters," *Med. Phys.* **21**, pp. 193–201, 8 1994.
9. W. Backfrieder, R. Baumgartner, M. Samal, E. Moser, and H. Bergmann, "Quantification of intensity variations in functional mr images using rotated principal components," *Phys. Med. Biol.* **41**, pp. 1425–1438, 8 1996.
10. M. McKeown, T. Jung, S. Makeig, G. Brown, T. Jung, S. Kindermann, A. Bell, and T. Sejnowski, "Spatially independent activity patterns in functional magnetic resonance imaging data during the stroop color-naming task," *Proc. Natl. Acad. Sci.* **95**, pp. 803–810, 8 1998.
11. M. McKeown, T. Jung, S. Makeig, G. Brown, T. Jung, S. Kindermann, A. Bell, and T. Sejnowski, "Analysis of fmri data by blind separation into independent spatial components," *Human Brain Mapping* **6**, pp. 160–188, 8 1998.
12. G. Scarth, M. McIntrye, B. Wowk, and R. Samorjai, "Detection novelty in functional imaging using fuzzy clustering," *Proc. SMR 3rd Annu. Meeting* **95**, pp. 238–242, 8 1995.
13. K. Chuang, M. Chiu, C. Lin, and J. Chen, "Model-free functional mri analysis using kohonen clustering neural network and fuzzy c-means," *IEEE Transaction on Medical Imaging* **18**, pp. 1117–1128, 8 1999.
14. R. Baumgartner, L. Ryder, W. Richter, R. Summers, M. Jarmasz, and R. Somorjai, "Comparison of two exploratory data analysis methods for fmri: fuzzy clustering versus principal component analysis," *Magnetic Resonance Imaging* **18**, pp. 89–94, 8 2000.
15. A. Wismüller, O. Lange, D. Dersch, G. Leinsinger, K. Hahn, B. Pütz, and D. Auer, "Cluster analysis of biomedical image time–series," *International Journal on Computer Vision* **46**, pp. 102–128, 2 2002.
16. H. Fisher and J. Hennig, "Clustering of functional mr data," *Proc. ISMRM 4rd Annu. Meeting* **96**, pp. 1179–1183, 8 1996.
17. S. Ngan and X. Hu, "Analysis of fmri imaging data using self-organizing mapping with spatial connectivity," *Magn. Reson. Med.* **41**, pp. 939–946, 8 1999.
18. T. Martinetz, S. Berkovich, and K. Schulten, "Neural gas network for vector quantization and its application to time-series prediction," *IEEE Transactions on Neural Networks* **4**, pp. 558–569, 4 1993.
19. F. Frisone, P. Vitali, G. Rodriguez, A. Pilot, F. Nobili, F. Sardanelli, and M. Rosa, "Self-organizing neural networks for fmri signal detetction," *5th International Conference on Functional Mapping of the Human Brain* **Poster No. 63**, 5 1999.
20. R. Woods, S. Cherry, and J. Mazziotta, "Rapid automated algorithm for aligning and reslicing pet images," *Journal of Computer Assisted Tomography* **16**, pp. 620–633, 8 1992.