

Model-Free Functional MRI Analysis Using Cluster-Based Methods

Thomas Dan Otto, Anke Meyer-Bäse,

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

Monica Hurdal, DeWitt Sumners,

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

Dorothee Auer, and

Max Planck Institute of Psychiatry, Munich 80804 Germany

Axel Wismüller

Department of Clinical Radiology, Ludwig-Maximilians University, Munich 80336 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 "neural gas" network 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 Kohonen's self-organizing map and with a minimal free energy vector quantizer is done in a systematic fMRI study showing comparative quantitative evaluations. The most important findings in this paper are: (1) the "neural gas" network outperforms the other two methods in terms of detecting small activation areas, and (2) computed reference functions reveal that the "neural gas" network outperforms the other two methods. The applicability of the new algorithm is demonstrated on experimental data.

Keywords: Minimal free energy vector quantization, "neural gas" network, self-organizing map, fMRI

1. INTRODUCTION

Functional magnetic resonance imaging with high temporal and spatial resolution is a potential method to map rapid and fine activation patterns of the human brain.¹⁻⁵ According to both theoretical estimations and experimental results,^{1,6,7} an activated signal variation appears very low on a clinical scanner. Thus, analysis methods are required to find the response waveforms and associated activated regions. Generally, these methods can be divided into two categories depending on whether or not they require prior knowledge about activation patterns: model-based and model-free. To eliminate the bias and limitation of model-based analysis methods and to satisfy the demand to analyze data with complicated experimental conditions, analysis methods that do not rely on any assumed model of functional response are necessary. 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.

The second method, fuzzy clustering analysis¹²⁻¹⁵ or self-organizing map,^{16,17,15} attempts to classify time signals of the brain into several patterns according to temporal similarity among these signals.

Email: amb@eng.fsu.edu

Paper-location-URL: <http://www.dtro.e-technik.th-darmstadt.de/amb/papers>

In this paper, we propose to employ the "neural gas" network to increase the analysis power without sacrificing 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) 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 than 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 Euclidian space which is subsequently partitioned into clusters based on the proximity of the input data.

2.1. 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

2.2. 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 in¹⁵ 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 explicitly given error function.

2.3. "NEURAL GAS" NETWORK

The "neural-gas" algorithm¹⁸ 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

$$\Delta \mathbf{w}_i = \epsilon e^{-k_i(\mathbf{w}, \mathbf{w}_i/\lambda)} (\mathbf{x} - \mathbf{w}_i) \quad i = 1, \dots, N \quad (1)$$

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.

In¹⁸ 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 in¹⁹ 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 stimulation periods with rest. Simulation and rest periods comprised 10 repetitions each, i.e. 30s. In one subject eight slices with 64 images (TR/TE=4000/66msec) were obtained starting with a period of rest lasting 10 repetitions (i.e. 40s) followed by three periods of stimulation alternating with periods of rest comprising 9 repetitions each, i.e. 36 s. 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) assessment of cluster assignment maps, (2) task-related activation maps and (3) associated time-courses. 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 three clustering techniques, SOM, "neural gas" network, and minimal free energy 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.

Therefore, it is significant to find a fixed number of CVs that can theoretically predict new observations under same conditions. To do so, we compared the three proposed algorithms for 8, 16, and 36 CVs in terms of: (1) the component time course most closely matching the visual task reference function which is found by each of the three techniques, (2) the normalized quantization error, and (3) the necessary simulation time.

The obtained results are plotted in Figure 1. In the following we will give the set parameters. For SOM we choose: (1) an one-dimensional lattice, and (2) the maximal number of iterations. 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. 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.

From Figure 1 we see that the lowest quantization error is achieved by the "neural gas" network, which at the same time provides the best correlation for eight CVs. It also outperforms the minimal free energy VQ. However, both can detect small regions.

The clustering results for the new method, the "neural gas" network, are shown in Figure 3 and 2. Figure 2 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 Figure 3 can be viewed as the cluster-specific weighted average of all pixel time courses.

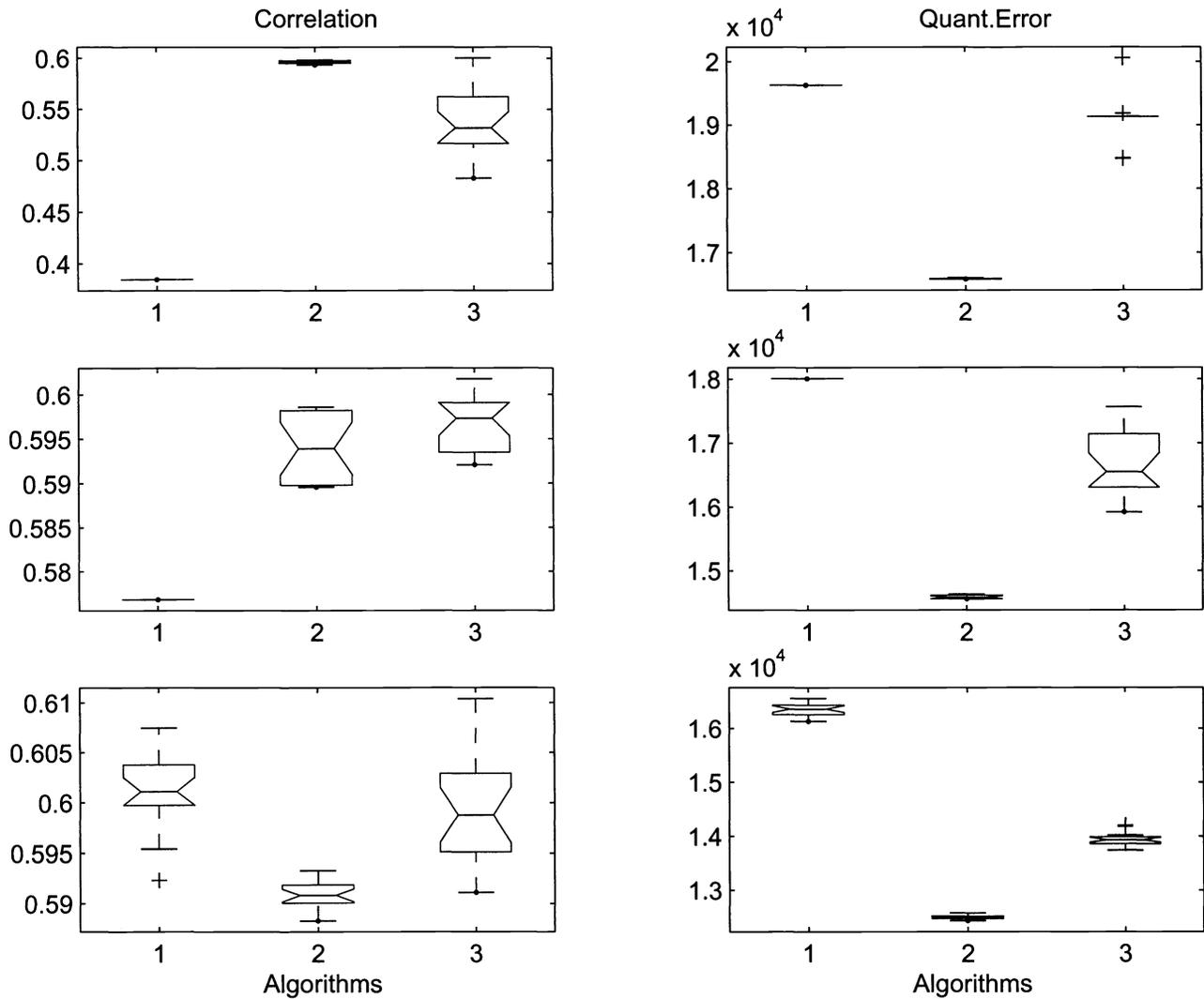


Figure 1. Estimation results for the optimal number of codebook vectors for the three clustering techniques, and considering 20 runs. From top to bottom, the results for 8, 16, and 36 CVs are visualized. The techniques 1 to 3 are: SOM, "neural gas" network, and minimal free energy VQ. The plots show the best found correlation coefficients between the found consistently task-related (CTR) components to the given task reference function, and the normalized quantization error.

3.2. CHARACTERIZATION OF TASK-RELATED EFFECTS

For all subjects, and runs, cluster assignment maps, unique task-related activation maps and associated time-courses were obtained by all three clustering techniques.

Figure 4 shows for 36 CVs the component time course most closely associated with the visual task for all three clustering techniques. The best results are achieved by the "neural gas" network and the minimal energy free VQ, yielding almost identical reference functions, and a correlation coefficient of $r = 0.9984$ between those two.

4. CONCLUSION

In the present paper, we have experimentally compared two proven clustering algorithms, the SOM and the minimal free energy vector quantization, with a less known algorithm, the "neural gas" network. The goal of the paper was

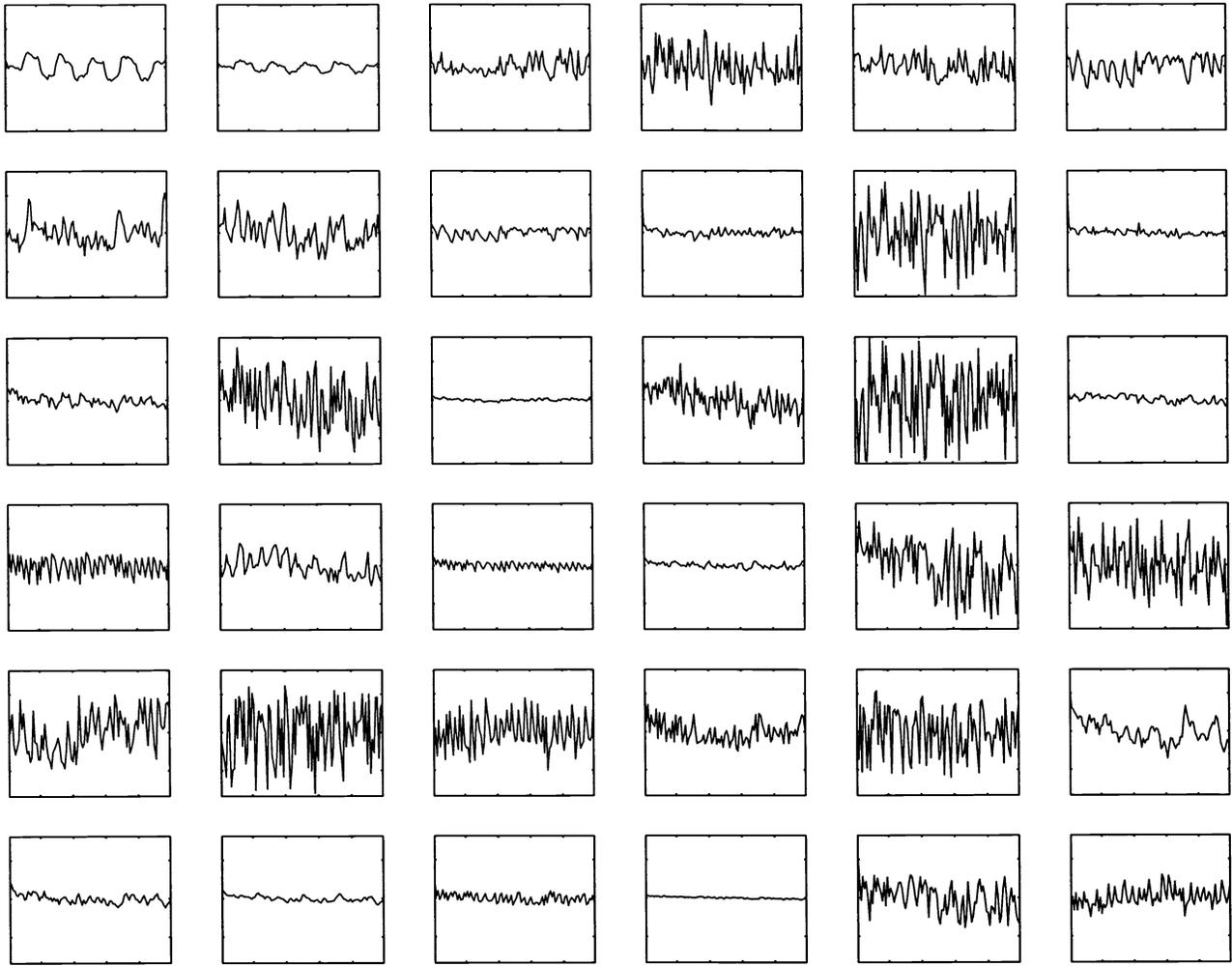


Figure 2. Cluster assignment maps for cluster analysis based on the "neural gas" network of a visual stimulation fMRI experiment and obtained for 36 CVs.

to demonstrate that unsupervised clustering techniques represent a 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 both "neural gas" network and minimal free energy VQ outperform in terms of found component time course, quantization error, and activation maps, SOM when applied to fMRI studies. However, SOM turned out to be the fastest algorithm, while the "neural gas" network proved to be the best clustering technique. 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.

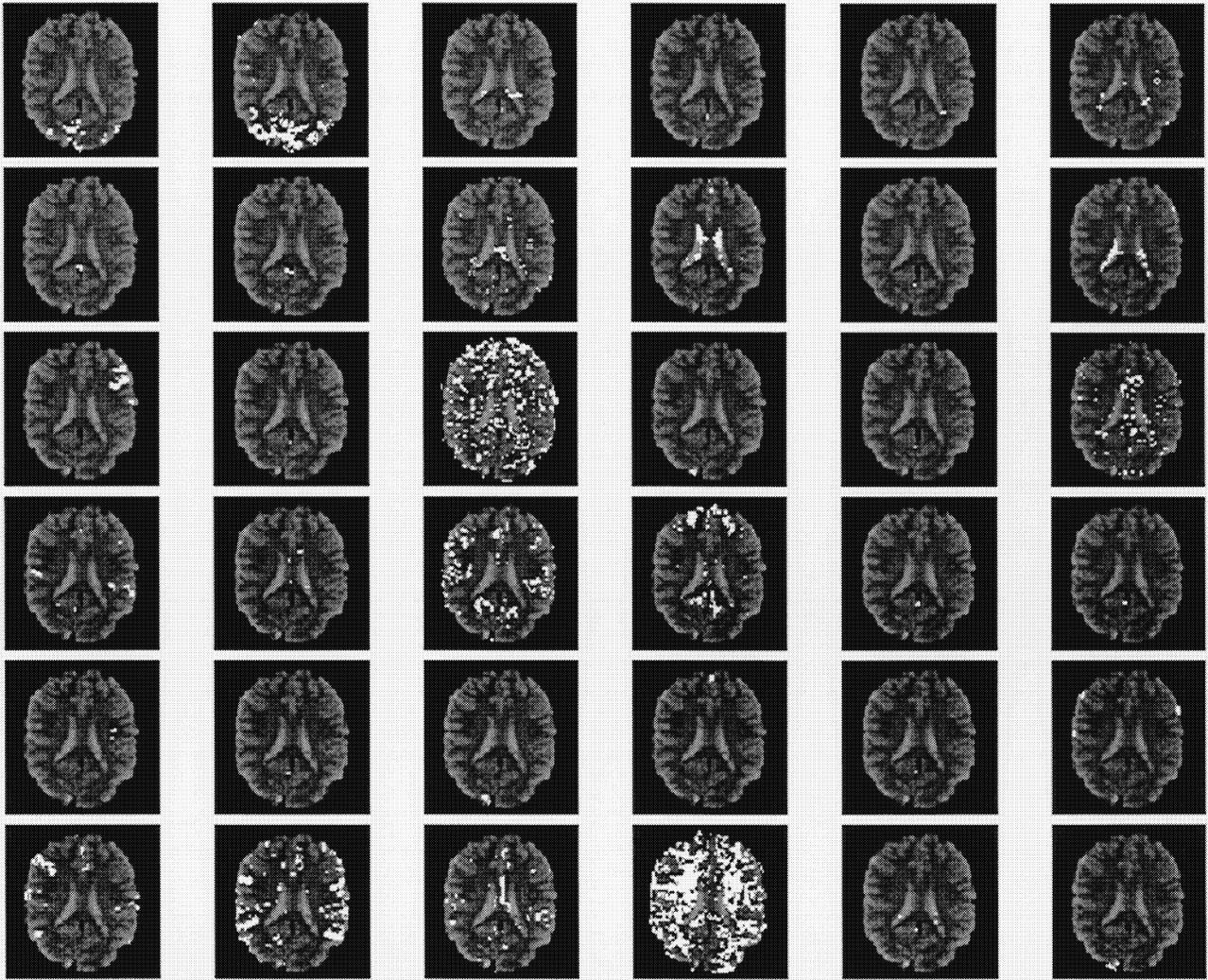


Figure 3. Associated codebook vectors for the "neural gas" network clustering as shown in Figure 2. The CVs were sorted according to the given reference function.

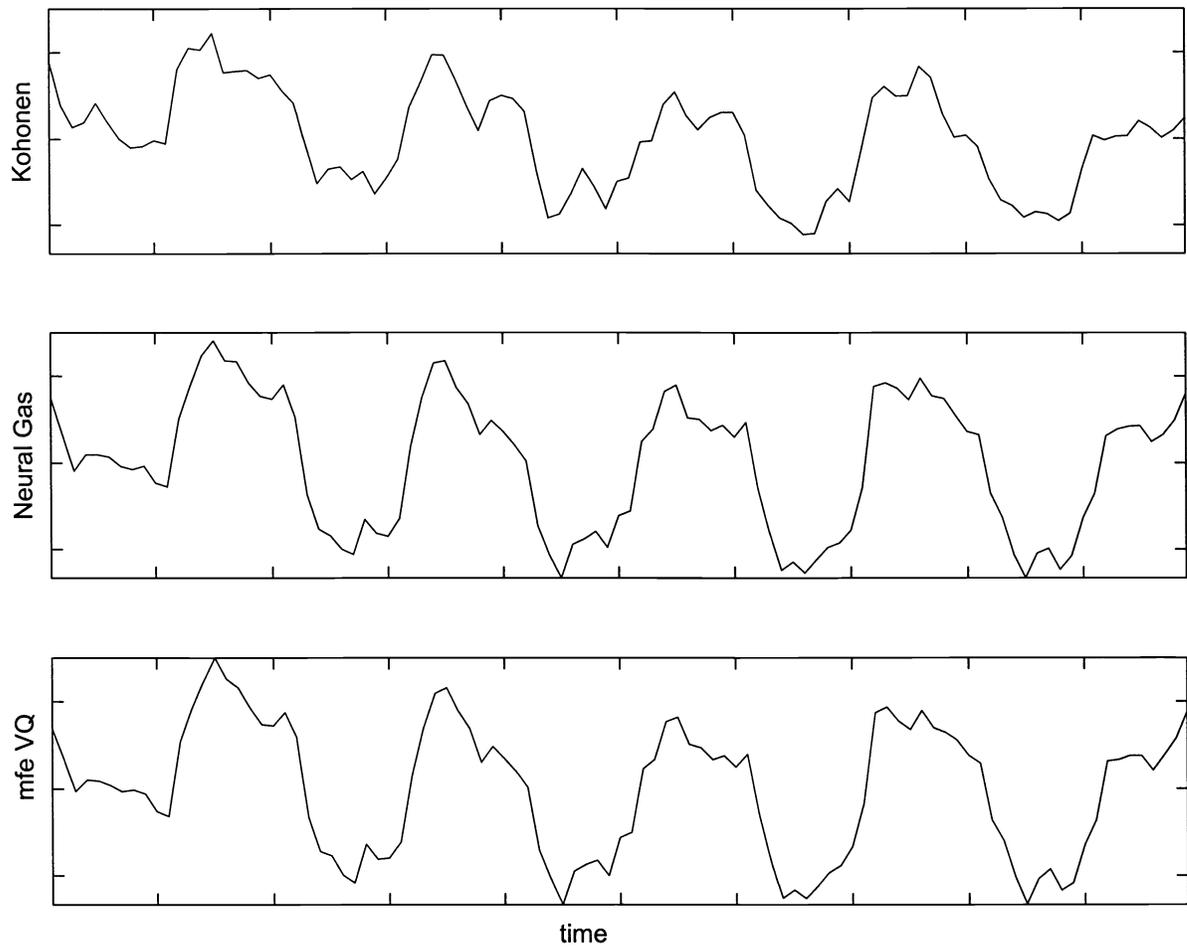


Figure 4. Computed reference functions for the three techniques, SOM, "neural gas" network, and minimal free energy VQ, for 36 code vectors. The found correlation coefficients are: $r = 0.51$ for SOM, $r = 0.56$ for "neural gas" network, and $r = 0.58$ for minimal free energy VQ.

REFERENCES

1. K. Kwong, J. Belliveau, and D. Chesler, "Dynamic magnetic resonance imaging of human brain activity during primary sensor stimulation," 8 1992.
2. S. Ogawa, D. Tank, and R. Menon, "Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging," 8 1992.
3. P. Bandettini, E. Wong, R. Hinks, R. Tikofski, and J. Hyde, "Time course epi of human brain function during task activation," 8 1992.
4. J. Frahm, K. Merboldt, and W. Hanicke, "Functional mri of human brain activation at high spatial resolution," 8 1992.
5. K. Kwong, "Functional magnetic resonance imaging with echo planar imaging," 8 1992.
6. J. Boxerman, P. Bandettini, K. Kwong, and J. Baker, "The intravascular contribution to fmri signal change: Monte carlo modeling and diffusion-weighted studies in vivo," 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," 8 1993.
8. J. Sychra, P. Bandettini, N. Bhattacharya, and Q. Lin, "Synthetic images by subspace transforms i. principal components images and related filters," 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," 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," 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," 8 1998.
12. G. Scarth, M. McIntyre, B. Wowk, and R. Samorjai, "Detection novelty in functional imaging using fuzzy clustering," 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," 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," 8 2000.
15. A. Wismuller, O. Lange, D. Dersch, G. Leinsinger, K. Hahn, B. Pütz, and D. Auer, "Cluster analysis of biomedical image time-series," 2 2002.
16. H. Fisher and J. Hennig, "Clustering of functional mr data," 8 1996.
17. S. Ngan and X. Hu, "Analysis of fmri imaging data using self-organizing mapping with spatial connectivity," 8 1999.
18. T. Martinetz, S. Berkovich, and K. Schulten, "Neural gas network for vector quantization and its application to time-series prediction," 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 detection," 5 1999.
20. R. Woods, S. Cherry, and J. Mazziotta, "Rapid automated algorithm for aligning and reslicing pet images," 8 1992.