

Fuzzy approach to incorporate hemodynamic variability and contextual information for detection of brain activation

Juan Zhou^a, Jagath C. Rajapakse^{a,b,*}

^a *Bioinformatics Research Center, School of Computer Engineering, Nanyang Technological University, Block N4, 50 Nanyang Avenue, Singapore 639798, Singapore*

^b *Singapore-MIT Alliance, Singapore*

ARTICLE INFO

Available online 1 July 2008

Keywords:

Brain activation
Contextual modeling
fMRI
Fuzzy *c*-means clustering
Feature extraction
Individual variability

ABSTRACT

We propose to use fuzzy *c*-means clustering with contextual modeling on features extracted from fMRI data for detection of brain activation. Five discriminating features are extracted from fMRI data by using a sequence of temporal-sliding-windows. Fuzzy membership maps of individual subjects obtained through clustering with spatial regularization is capable of taking into account both hemodynamic variability and contextual information of brain activation. The present method outperforms statistical parametric mapping (SPM) approach on experiments with synthetic fMRI data contaminated by both independent and correlated noise. Performance on real fMRI data are comparable to those obtained with SPM.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Functional magnetic resonance imaging (fMRI) is a non-invasive imaging modality measuring functional activity of the brain *in vivo* both spatially and temporally. fMRI signal results due to the changes of blood-oxygenation-level-dependent (BOLD) contrast, caused by the increase in blood oxygenation following neuronal activity. Detection of the changes of fMRI signal is non-trivial as BOLD signal change due to an input stimuli is very subtle, ranging from 1–5% on a 1.5 T scanner [20]. Furthermore, various noise and artifacts such as motion, electronic, physical, and physiological processes, significantly confound fMRI signal. Therefore, techniques analyzing fMRI signals should be insensitive to uncertainties and randomness of interference signals.

Methods to detect activated voxels from fMRI data fall into two categories: hypothesis-driven and data-driven methods. Statistical parametric mapping (SPM) [11] is a widely used hypothesis-driven method assuming a general linear model for fMR signal with a specific noise structure. It is voxel-based and tests the hypothesis about fMR time-series response on the stimuli by construction and assessment of spatially extended statistical processes based on Gaussian random fields (GRF). However, the actual relationship between the change of fMR signal and the stimuli presentation is nonlinear [29]; and the hemodynamic

response function (HRF) varies spatially and among subjects [24,33]. Moreover, the structure of noise in fMRI is not well understood and remains a contentious subject [7].

Data-driven methods do not make any assumptions on hemodynamic response and are considered more appropriate and powerful for fMRI analysis, especially when unknown or complex differential responses are expected [18]. Data-driven approaches can be broadly classified into transformation-based or clustering-based methods. Principle component analysis (PCA) [3] and independent component analysis (ICA) [2] transform original high-dimensional fMRI data into a low-dimensional space to separate brain activation and various noise sources. The ICA enables recovery of underlying task-related signals from other components such as artifacts and noise by decomposing fMRI data spatially [16] and temporally [6] in an exploratory manner or with stimulus as constraints [14,15]. Clustering techniques, such as self-organizing maps [13,19] and fuzzy clustering [5,9], attempt to classify fMR time-series of the brain into several patterns according to temporal similarity. Data-driven methods usually interpret the contents of one class or component as activations but how signals are divided into classes is difficult to ascertain or comprehend; a few classes related to activation could have physiological interpretation but interpretations of others are unknown. Other data-driven techniques for fMRI analysis include multi-resolution methods such as wavelet analysis [12].

Besides the activation measured at each brain voxel, fMRI carries contextual information as neighboring voxels often have similar characteristics and belong to the same class. Gaussian smoothing is often applied to enhance signal to noise ratio (SNR) before statistical analysis, accounting for spatial dependency

* Corresponding author at: Bioinformatics Research Center, School of Computer Engineering, Nanyang Technological University, Block N4, 50 Nanyang Avenue, Singapore 639798, Singapore. Tel.: +65 6790 5802; fax: +65 6792 6559.

E-mail address: asjagath@ntu.edu.sg (J.C. Rajapakse).

implicitly. This leads to overly smoothed images and a loss of high frequency information. Markov random fields (MRF) [27,25,10] and conditional random fields (CRF) [30] have been attempted to incorporate spatial and temporal correlations explicitly for the detection of brain activation. MRF based on mixture models with spatial regularization has been proposed for fMRI segmentation [32]. Autoregressive spatiotemporal models has also been proposed to incorporate tissue-type related noise priors as spatial constraints in fMRI analysis [33]. The above approaches still suffer from the assumptions made on structures of HRF and noise, the validity of such models depends on the extent to which data satisfy underlying assumptions.

Fuzzy c -means clustering (FCM) has been widely used in image segmentation, which incorporates fuzziness for the belongingness of each voxel to particular class or object. Contextual modeling is handled by a contextual regularized term into the cost function of FCM: for example, smooth constraints on the bias field of intensity inhomogeneity correction [21] and a term for immediate neighborhood information for tissue segmentation of anatomical MR images [1]. Kernelized version of FCM with spatial constraints has been recently proposed to further enhance the segmentation [8].

The present work is motivated by the need for a technique that handles the individual variability of hemodynamic responses and is less vulnerable to noise. In order to achieve this, we extract fuzzy features from fMR time-series and use unsupervised FCM with contextual modeling to identify activated voxels. To adapt to HRF variability across subjects, FCM is applied to each subject separately assuming no fixed structure for signals and noise. The cost function of FCM, regularized by spatial context, is adopted to capture local and neighboring information. Extracting fuzzy features from time-series reduces computational complexity and renders the ability to handle noise. FCM with contextual modeling on novel fuzzy feature space is similar to kernelized FCM on original fMR time-series but the advantage of our approach is that the kernel is well defined by the new feature space compared to normal Gaussian or polynomial function. The details of our approach are described in the next section. In the experiment and result section, the performance of the present approach is illustrated with functional activation detected on individual and group study on synthetic as well as on real fMRI data.

2. Method

Our method consists of two steps: (1) extracting fuzzy features from raw fMR time-series by temporal-sliding-windows (TSW); (2) FCM clustering with contextual constraints on the feature space.

2.1. Extracting features

Different brain voxels have different hemodynamic characteristics, for example, the time to reach the peak intensity of an activated voxels. We move a sequence of TSW over the time-series of a voxel to derive fuzzy features discriminating activated and non-activated voxels under each experimental condition. These features are derived independent of shape, magnitude, and delay of HRF. Let $\Psi : \Omega \times \Theta \rightarrow Y$ be a functional MR image where $\Omega \subset N^3$ denotes three-dimensional spatial domain of brain voxels, $\Theta = \{1, 2, \dots, n\}$ indexes n number of 3D scans taken during the experiment. Let $Y = \{y_{i,t} : i \in \Omega, t \in \Theta, y_{i,t} \in Q\}$ be the 4D fMRI data where Q denotes the range of image intensities and $y_{i,t}$ denotes the intensity of voxel i at time t .

Consider an experiment with only one condition, denoted by X , for notational simplicity; the technique of feature extraction is easily extended for fMR experiments involving several conditions

[37]. The condition X is presented with the reference (resting) state alternatively for P times in a single run while n 3D brain scans are taken. Each block of condition X is denoted by B_p , $p = 1, 2, \dots, P$. The block B_p lasts for a duration of length l_p from the beginning time denoted as b_p . The above notation represents a general paradigm design which applies to both block and event-related designs. A sequence of TSW for the condition X is constructed from fMR time-series as follows:

- (1) Create a sequence of P number of windows denoted by $W = \{W_p : p = 1, 2, \dots, P\}$, one window W_p for each condition block B_p . The length of window W_p is denoted by w_p , and let $w_p = l_p$. The initial starting point of window W_p is thus given by b_p .
- (2) Shift the sequence of windows W temporally forward by a sliding time interval s simultaneously, resulting in a new sequence of windows denoted by $W(s) = \{W_p(s) : s = 0, 1, \dots, S, p = 1, 2, \dots, P\}$. Depending on different inter-scan time, the maximum sliding time interval S varies: $S = 32/RT$ (seconds) based on the fact that the total length of HRF is approximately 32 s. Thus, the starting and ending time of window $W_p(s)$ is $b_p + s$ and $b_p + s + w_p - 1$; we denote them by $T_{p,1}(s)$ and $T_{p,2}(s)$, respectively, for notational simplicity.
- (3) Calculate average intensity $A_p(i, s)$ of each voxel i for each window $W_p(s)$, $s = 0, 1, \dots, S$ as

$$A_p(i, s) = \frac{\sum_{t=T_{p,1}(s)}^{\tilde{T}_{p,2}(s)} y_{i,t}}{\tilde{T}_{p,2}(s) - T_{p,1}(s)} \quad (1)$$

$$\text{where } \tilde{T}_{p,2}(s) = \min\{n, T_{p,2}(s)\}.$$

Thus, we observe a curve $A_p(i) = \{A_p(i, s) : s = 0, 1, \dots, S\}$ for each voxel i of each block B_p , whose shape is highly discriminative between activated and non-activated voxels. We refer it as *quasi-hemodynamic curve* (QHC) because it represents the HRF derived from the time-series in a data-driven manner. Five fuzzy discriminating features $F_k^p(i)$, $k = 1, 2, \dots, 5$, are extracted from QHC for each block B_p at each voxel i as follows:

- (1) Area under curve ratio for QHC:

$$F_1^p(i) = \frac{\sum_{s=0}^{\tilde{w}_p} A_p(i, s)}{(\max_s A_p(i, s) - \min_s A_p(i, s)) \cdot \tilde{w}_p} \quad (2)$$

$$\text{where } \tilde{w}_p = \min\{w_p, S\}.$$

- (2) Area difference ratio for QHC:

$$F_2^p(i) = \frac{\sum_{s=0}^{\tilde{w}_p} A_p(i, s)}{\sum_{s=\tilde{w}_p+1}^S A_p(i, s)} \quad (3)$$

- (3) Correlation between QHC $A_p(i)$ and the standard QHC, SA_p :

$$F_3^p(i) = \frac{\sum_{s=0}^{\tilde{w}_p} (A_p(i, s) - \bar{A}_p(i, s))(SA_p - \bar{SA}_p)}{\sqrt{\sum_{s=0}^{\tilde{w}_p} (A_p(i, s) - \bar{A}_p(i, s))^2 \sum_{s=0}^{\tilde{w}_p} (SA_p - \bar{SA}_p)^2}} \quad (4)$$

$$\text{where } SA_p = \{SA_p(s) = -(s - \tilde{w}_p/2)^2 : s = 0, 1, \dots, \tilde{w}_p\}.$$

- (4) Time ratio at peak amplitudes of QHC:

$$F_4^p(i) = \arg \max_{s \in [0, \tilde{w}_p]} A_p(i, s) / \tilde{w}_p \quad (5)$$

- (5) Time ratio at lowest amplitude for QHC:

$$F_5^p(i) = \arg \min_{s \in [0, \tilde{w}_p]} A_p(i, s) / \tilde{w}_p \quad (6)$$

Two QHC were normalized to within $[0, 1]$ before correlation computation in feature 3 for easy comparison among voxels. Since the shapes of QHC of activated and non-activated voxels

are usually different as seen later, the above five features could be significantly discriminating.

The above five features of each block are sum up over all blocks as

$$F_k(i) = \sum_{p=1}^P F_k^p(i)/P, \quad (7)$$

which leads to a robust 5D feature space $\{F_k(i) : k = 1, 2, \dots, 5; i \in \Omega\}$ at each voxel, which fuzzy features are less vulnerable to noise or changes in hemodynamic response. Note that we assume the duration of each block $l_p \leq S$ in the above definitions. For cases like $l_p > S$, we should apply the opposite settings, i.e., use window length $w_p = 32/RT$ and the maximum sliding time for each window $S_p = l_p$. The properties of the resulting curve is similar to QHC and the same features could be extracted.

2.2. FCM clustering with contextual constraints

Based on feature space developed in the previous section, we use FCM with spatial constraints to classify voxels into activated and non-activated classes. This scheme is able to adapt to the hemodynamic variability across subjects when unsupervised FCM is applied on each subject separately. FCM is able to provide (1) the strength of activation of each voxel to each condition, (2) a clear classification of voxels by comparing membership values, and (3) a rule-base for interpretation of activation patterns. For a given condition X , each voxel $i \in \Omega$ has a corresponding vector of five features $F(i) = \{F_k(i) : k = 1, 2, \dots, 5\}$. We use FCM on these features to classify voxels into two classes: activated or non-activated class.

The objective function of standard FCM partitioning the feature set $\{F(i) : i \in \Omega\}$ into C clusters is given by

$$J_m = \sum_{c=1}^C \sum_{i \in \Omega} u(c, i)^m \|F(i) - V(c)\|^2 \quad (8)$$

subject to constraints:

$$u(c, i) \in [0, 1], \quad \sum_{c=1}^C u(c, i) = 1, \quad \forall i \quad \text{and} \\ 0 < \sum_{i \in \Omega} u(c, i) < |\Omega|, \quad \forall c \quad (9)$$

where $u(c, i)$ is the fuzzy membership value of voxel i belonging to class c , $V(c)$ is the feature vector of the centroid of class c , $\|\cdot\|$ stands for Euclidean norm, and parameter m is the weighting exponent of fuzzy memberships.

Standard FCM fails to segment images corrupted by noise, outliers, and other artifacts [8]. The robustness of FCM to noise can be increased by directly modifying the objective function in Eq. (8) to incorporate contextual information [1]:

$$J_m = \sum_{c=1}^C \sum_{i \in \Omega} u(c, i)^m \|F(i) - V(c)\|^2 \\ + \frac{\alpha}{|N(i)|} \sum_{c=1}^C \sum_{i \in \Omega} u(c, i)^m \sum_{r \in N(i)} \|F(r) - V(c)\|^2 \quad (10)$$

where $N(i)$ stands for the set of neighbors around voxels i and $|N(i)|$ is the number of valid voxels in the defined neighborhood. The parameter α controls the effect of regularization from neighboring voxels. The clusters are obtained by solving

$$\min J_m \quad \text{s.t.} \quad \sum_{c=1}^C u(c, i) = 1, \quad \forall i \quad (11)$$

With this minimization, high membership values are given to voxels nearer to the centroid and low membership values are given to those far away from the centroid. The Lagrange multiplier is adopted to include the constraints into optimization and the augmented object function becomes

$$J'_m = J_m + \lambda \left(1 - \sum_{c=1}^C u(c, i) \right) \quad (12)$$

This problem can be solved by taking the derivatives of J'_m with $u(c, i)$ and applying the constraints. The necessary conditions for the local minimum are

$$u(c, i) = \frac{\left(\|F(i) - V(c)\|^2 + \frac{\alpha}{|N(i)|} \sum_{r \in N(i)} \|F(r) - V(c)\|^2 \right)^{-1/(m-1)}}{\sum_{c=1}^C \left(\|F(i) - V(c)\|^2 + \frac{\alpha}{|N(i)|} \sum_{r \in N(i)} \|F(r) - V(c)\|^2 \right)^{-1/(m-1)}} \quad (13)$$

$$V(c) = \frac{\sum_{i \in \Omega} u(c, i)^m \left(F(i) + \frac{\alpha}{|N(i)|} \sum_{r \in N(i)} F(r) \right)}{(1 + \alpha) \sum_{i \in \Omega} u(c, i)^m} \quad (14)$$

FCM clustering with contextual modeling can be achieved in the following iterative updating procedure:

- (1) For a fixed number of C , set initial cluster centroid $V(c)$ and choose $\varepsilon > 0$ to a very small value.
- (2) Calculate the fuzzy membership values of each voxel to all clusters by Eq. (13).
- (3) Update the centroids based on new membership values by Eq. (14).
- (3) Repeat steps 2 and 3 until the average absolute difference of centroids obtained in consecutive rounds are less than ε .

The final fuzzy membership map for both activated and non-activated voxels is able to provide a concrete segmentation of activated and non-activated brain regions from fMRI. The voxel is assigned to the class with highest fuzzy membership value.

3. Experiments and results

We demonstrate our approach on both synthetic and real fMRI data. A comparison between the results produced by our FCM clustering approach and SPM is given.

3.1. Synthetic data

A 2D synthetic functional dataset consisting of six cycles (with eight ON and eight OFF scans, $TR = 2$ s, $n = 96$) was simulated. The response of activated voxels was generated by convolving a box-car time-series with HRF, using a mixture of two gamma functions, while non-activated voxels were kept constant at zero amplitude. The ground truth of activation pattern is shown in Fig. 1a. Synthetic images with different levels of independent noise (Gaussian) and spatially correlated noise (by averaging the neighboring i.i.d. Gaussian noise) were tested. Five synthetic time-series were simulated based on different HRF by varying its parameters: the delay of response and undershoot relative to onset, the dispersion of response and undershoot, ratio of response to undershoot, and the total length of HRF function. Fig. 1b illustrates HRF function at different parameter values. The purpose is to test the vulnerability of FCM and SPM to the variability of HRF across subjects and their robustness against noise.

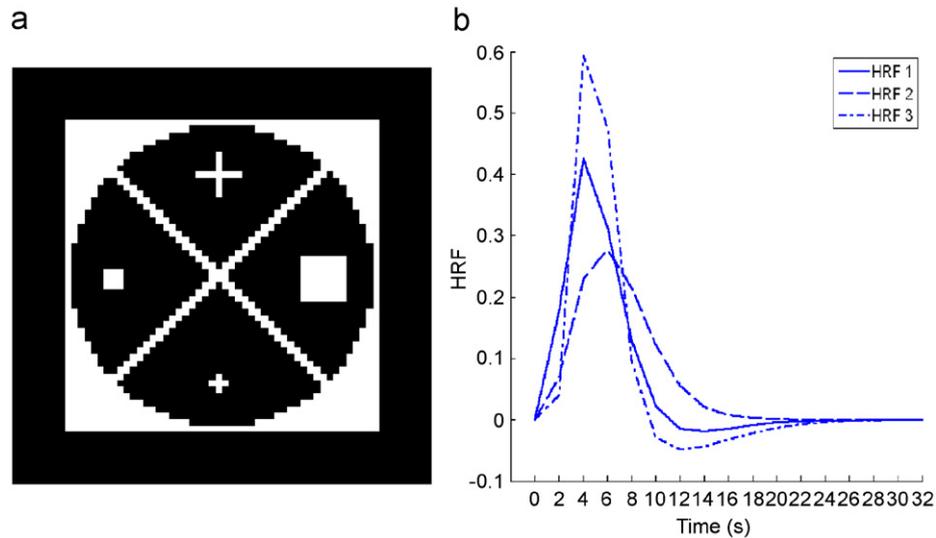


Fig. 1. Generation of synthetic fMRI data: (a) true activation region; and (b) example of three HRFs of the synthetic time-series of different subjects.

SPM analysis for both real and synthetic datasets used the standard procedure implemented in SPM2 [31]. A 6 mm 3D Gaussian filter was applied to increase the SNR before statistical analysis. Canonical hemodynamic response was used as basis function for statistical modeling. Confounding effects of fluctuations of global mean were removed by proportional scaling, and low-frequency noise was removed with a high pass temporal filter (128 s) applied to the fMRI time-series at each voxel. Specific effects were tested by applying appropriate linear contrasts and t -statistical parametric maps were used to assess significant hemodynamic changes. We report activations in voxels below a threshold of $p < 0.05$, which were corrected for multiple comparisons using false discovery rate (FDR) [4].

FCM was performed on original (unsmoothed) time-series. Features were first extracted using a sequence of TSW of width $w_p = 8$. The parameter settings were: $\alpha = 3$, $m = 2$, $\varepsilon = 0.001$. We empirically determined the optimal value of α . A six-voxel neighborhood was used for contextual modeling of the target voxel. Initial cluster centroids were set to the vector of voxels with the highest and lowest correlation with input stimulus convolved with HRF. Extracted features were individually scaled to $[0, 1]$ before applying FCM such that each feature plays an equal weight.

Various levels of noise in both independent and correlated cases were tested: $\text{SNR} = \{2.0, 1.2, 0.45\}$. The performance of FCM for functional activation detection on a group of five synthetic fMRI time-series was compared to SPM and our previous work fuzzy feature modeling (FFM) [37] by plotting the ROC curves as shown in Fig. 2. As seen, present FCM approach outperforms SPM for data with both independent and correlated noise, especially at high level of correlated noise frequently embedded inside real fMRI data. The previous FFM approach does not incorporate contextual information and thus suffers from high level of independent noise, although it performs better than contextual FCM in other cases. It is suspected that the fuzzy feature model built in FFM is able to classify nonlinear functions by incremental learning while current version of FCM, though can incorporate contextual information, may lack the ability of adaptive learning. On the other hand, thresholding is required in FFM whereas contextual FCM automatically identify the activated and non-activated voxels by winner-takes-all rule.

The final segmentation produced by FCM and SPM (t -contrast with $\text{FDR } p < 0.05$) is shown in Fig. 3. As seen, our FCM approach is able to discover more accurate and detailed activation than SPM,

especially in high correlated noise case. The elegance of FCM is that no thresholding is involved; activated voxels are automatically identified by comparing the fuzzy membership values of the classes, i.e., with winner-takes-all. For SPM, thresholding methods should be catered to correct for multiple comparisons. Family-wise-error (FWE) thresholding method usually produces more conserved results than false-discovery-rate (FDR) method. In this experiment, we found that FDR is more appropriate for images with correlated noise while FWE is more appropriate for images with independent noise. This is also seen by thresholded images in Figs. 3 and 4. Here, the results thresholded by FDR are reported. SPM uses a fixed HRF and therefore cannot adapt to HRF variability. Experiments showed that SPM suffered more from correlated noise than Gaussian noise whereas contextual FCM performs much better under the case of correlated noise. Overly smoothing in SPM leads to loss of high-frequency information and the assumptions of HRF and noise prevent it from detecting highly variable activation. Moreover, the resulting centroids of activated and non-activated classes by contextual FCM correspond well to various parameters of HRF during generation of synthetic fMRI data, which means that the extracted features are capable of capturing the properties of HRF.

3.2. Real data

3.2.1. Visual task

A set of real fMRI data obtained from experiments with a visual task were analyzed, see Rajapakse et al. [24] for further details about this data. For SPM analysis, all functional images were first corrected for movement artifacts, resampled, and smoothed with a 3D Gaussian filter having $\text{FWHM} = 4.47$ mm. T -contrast is used for statistical analysis in SPM2 with canonical HRF as basis function. Voxels with $p < 0.05$ corrected using FWE is determined to be activated. For FCM approach, the same parameter settings were used as for synthetic data on original unsmoothed images. Features are first extracted using a sequence of TSW of width $w_p = 4$.

In order to illustrate the usefulness of the extracted 5D feature space, QHC were derived. Although QHC varies across subjects, brain regions and tasks, the characteristics of QHC often comprises similar discriminating features. This is illustrated in Fig. 5 showing typical QHC for an activated voxel (top) and

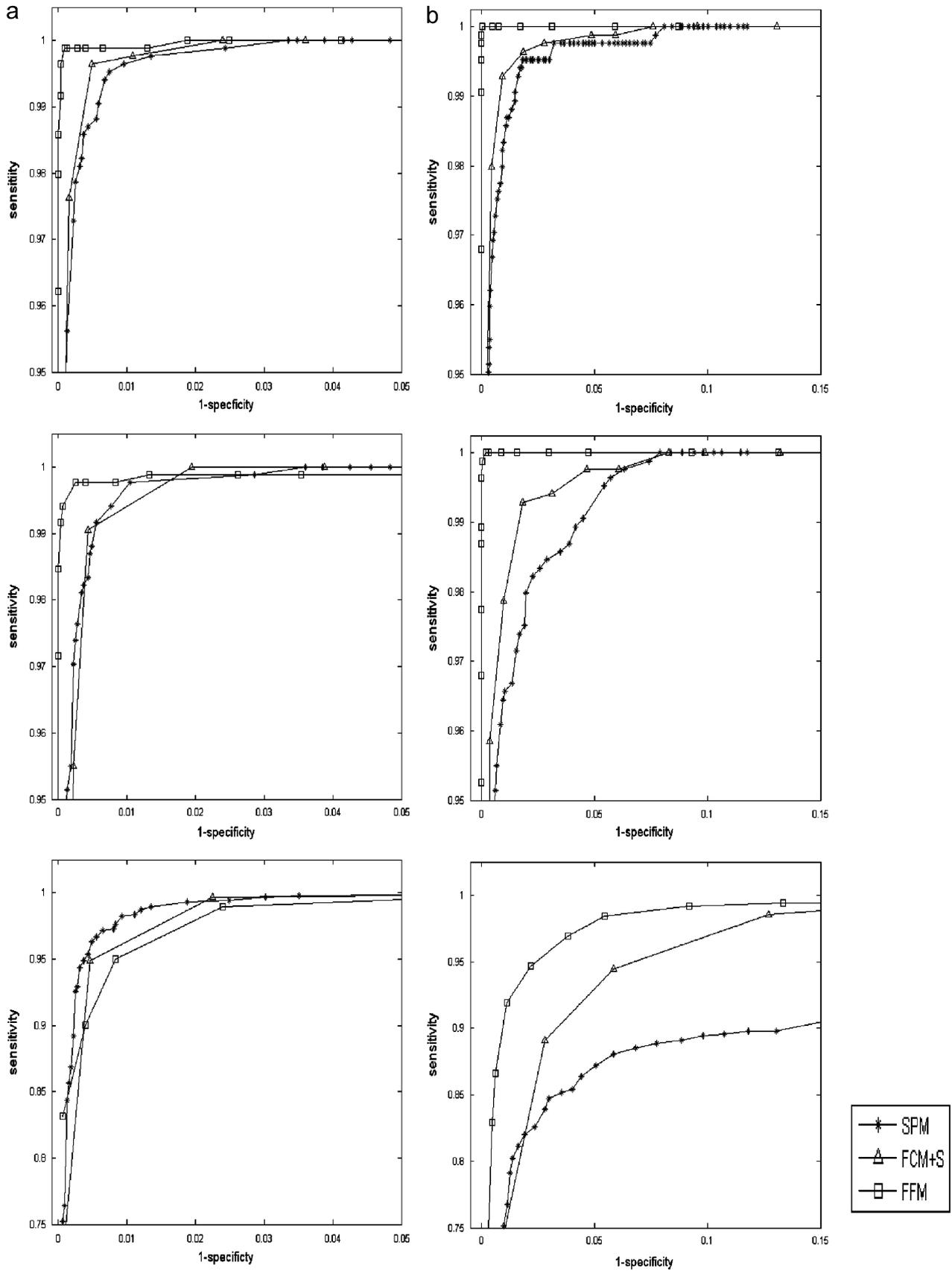


Fig. 2. ROC curves for detecting activation on a group study of five synthetic time-series with different HRF by SPM, FFM, and FCM methods. Different levels of (a) independent and (b) correlated noise were tested: row 1 SNR = 2.0, row 2 SNR = 1.2, and row 3 SNR = 0.45.

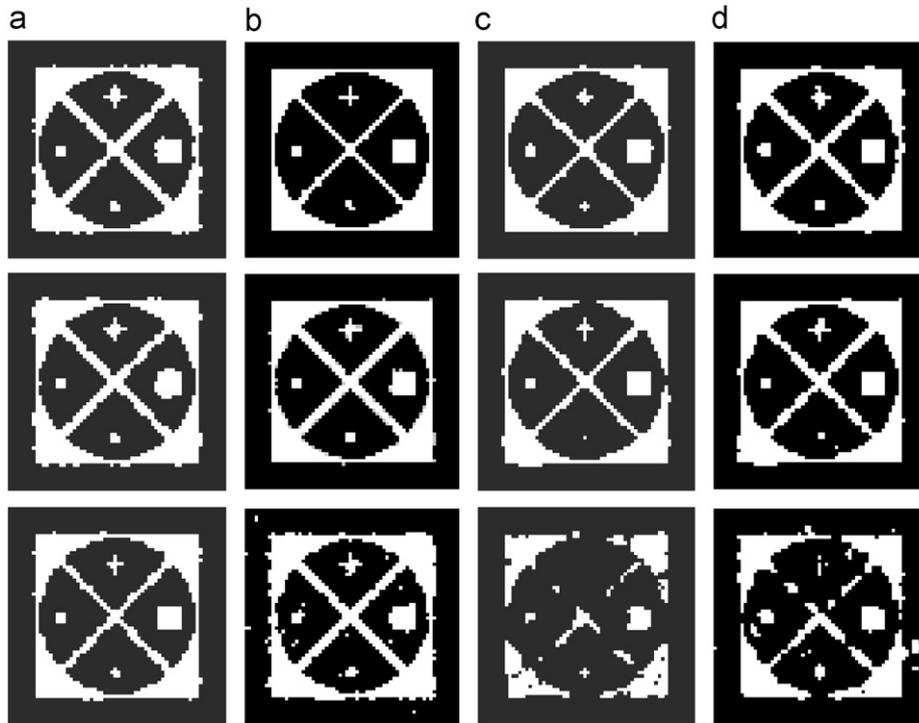


Fig. 3. Comparison of detected activation by SPM and FCM from a group study of synthetic data at different levels of independent noise (a) and (b) and correlated noise (c) and (d). Row 1: SNR = 2.0, row 2: SNR = 1.2, and row 3: SNR = 0.45.

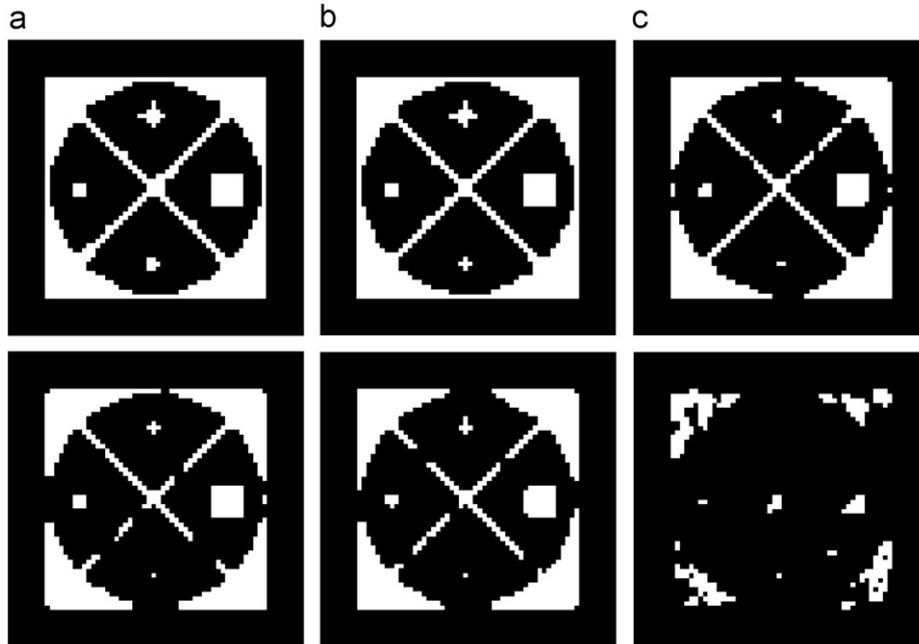


Fig. 4. The detected activation by thresholding with family-wise-error (FWE) in SPM from a group study of synthetic data at different levels of independent noise (first row) and correlated noise (second row). (a) SNR = 2.0, (b) SNR = 1.2, and (c) SNR = 0.45.

non-activated voxel (bottom) in real fMRI data of visual (a) and motor (b) task, respectively. Despite the differences between QHC of two tasks, evidently, the activated and non-activated voxels have quite different QHC shapes of brain regions and hence common discriminating features could still be discovered with lower degrees of uncertainty.

Fig. 6 shows the detected activated regions for visual task on three axial slices of one subject by both SPM and FCM approaches. Since the ground truth of brain activation is unknown, it is difficult to compare activation patterns quantitatively, but still activation was found in expected regions of visual cortex for both approaches. In addition, the activation found by FCM exactly

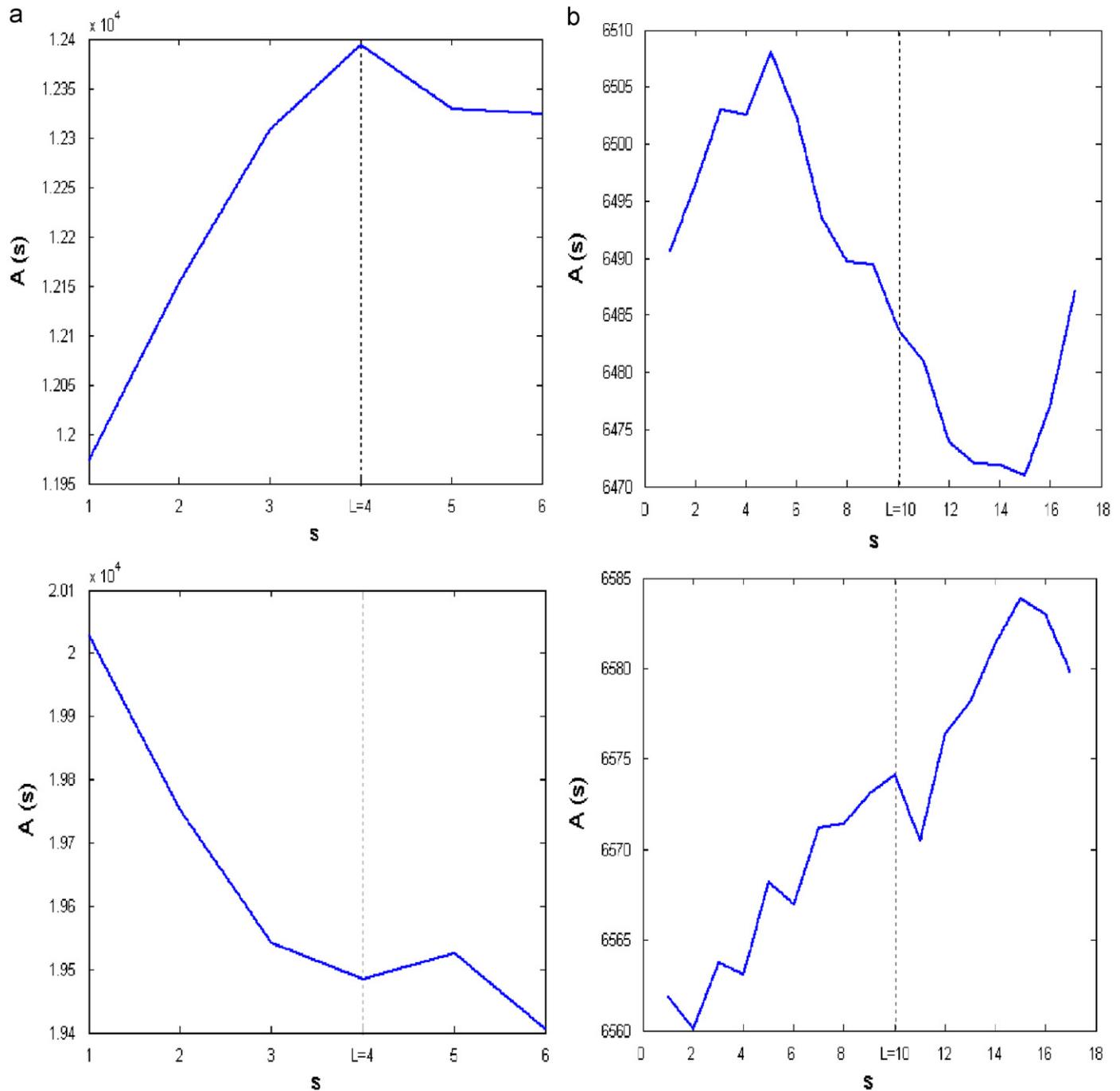


Fig. 5. QHC extracted from fMRI data of visual (a) and motor (b) tasks for activated (top) and non-activated (bottom) voxels, respectively.

follows the curvature of gray tissue matter as seen while the activation produced by SPM does not.

3.2.2. Counting Stroop task

fMRI data in a counting Stroop task investigating the performance of females with fragile X-syndrome on the cognitive interference processing compared to normal subjects were analyzed. There are two conditions in the task: neutral and interference, see Tamm et al. [28] for details about the task and data collection. Standard SPM analysis was performed based on HRF with time and derivatives and thresholded by FDR ($p < 0.05$). The same parameter settings used in contextual FCM approach on

synthetic data experiments were used on this data. Features are first extracted using a sequence of TSW of width equals to 15. A female patient is examined for activated regions in both neutral and interference counting Stroop condition. Fig. 7 shows the results by SPM and FCM for each condition, respectively. As seen, their results are quite similar. The findings are consistent with previous literature [28,35], unique activation were found in left inferior/middle frontal gyrus (BA 45, 46), left supplementary motor area (BA 6) and right middle/inferior frontal gyrus (BA 9/47) for interference condition, while left and right putamen, left hippocampus, left parahippocampal gyrus, right superior temporal gyrus (BA 22) and right posterior insula were activated in neutral condition.

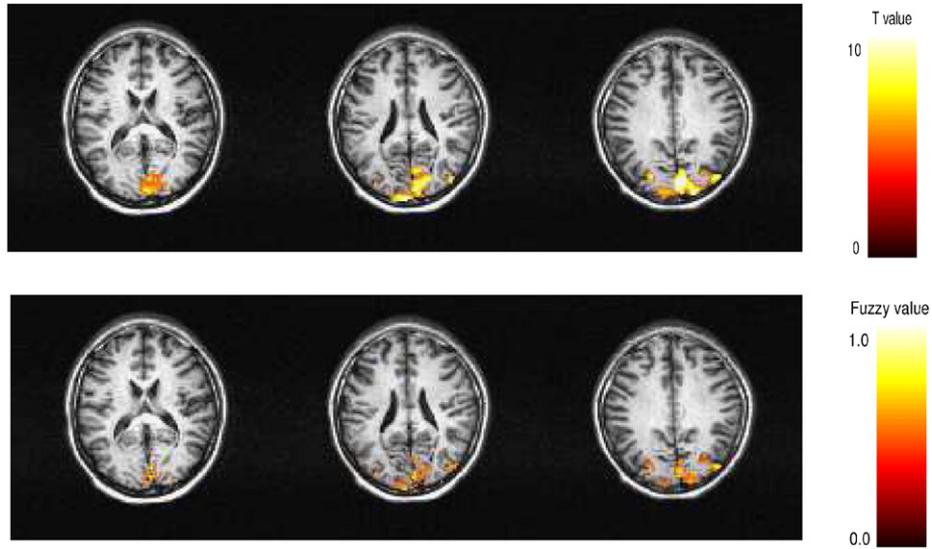


Fig. 6. Detected activation on three axial slices by SPM (top) and FCM (bottom) for the visual task.

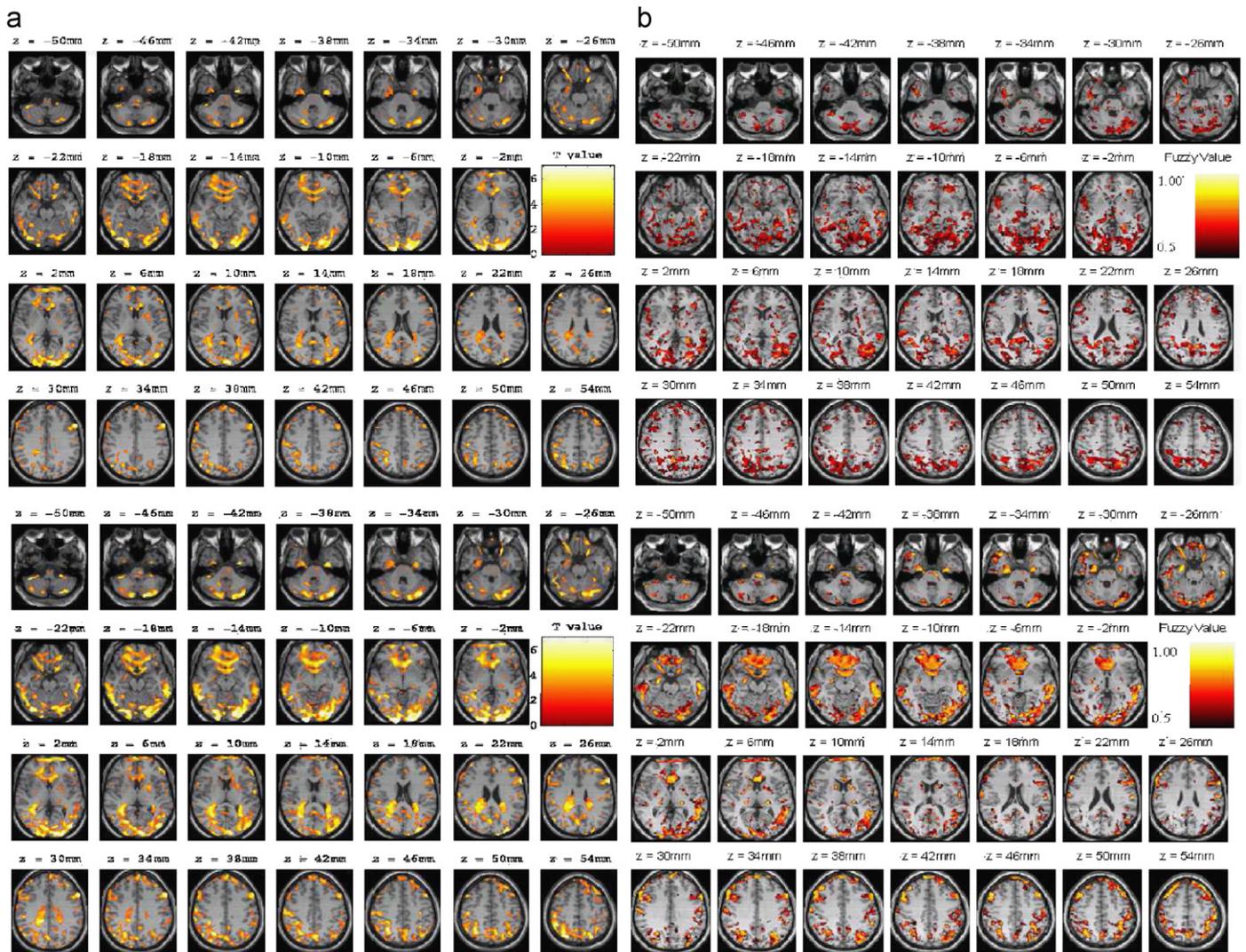


Fig. 7. Detected activation on selected axial slices by (a) SPM and (b) FCM for counting Stroop task on a female X Syndrome patient: neutral condition (top) and interference condition (bottom).

4. Conclusion

Using a novel fuzzy feature extraction method, we are able to convert 4D fMRI dataset into a simpler and robust feature space for detection of functional activation. Unsupervised fuzzy *c*-means clustering with contextual modeling was applied to the feature space to detect brain activation. Experiments on both synthetic and real fMRI data showed that our FCM approach is less vulnerable to both independent and correlated noise incorporating neighborhood information of brain activation. Activated and non-activated voxels for the condition of interest were discovered simultaneously and explicitly. A group study on synthetic fMRI data generated by different HRF further illustrates that FCM is capable of handling HRF variability across subjects. Our future work will include incorporating the variability of brain tissues [23] and structures [36] into activation detection by multi-modality image fusion [22].

References

- [1] M.N. Ahmed, S.M. Yamany, N. Mohamed, A.A. Farag, T. Moriarty, A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data, *IEEE Trans. Med. Imaging* 21 (3) (2002) 193–199.
- [2] S. Amari, A. Cichocki, H.H. Yang, A new learning algorithm for blind signal separation, *Adv. Neural Inform. Process. Syst.* 8 (1996) 752–763.
- [3] A.H. Andersen, D.M. Gash, M.J. Avison, Principal component analysis of the dynamic response measured by fMRI: a generalized linear systems framework, *Magn. Reson. Imaging* 17 (6) (1999) 795–815.
- [4] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, *J. R. Statist. Soc. Ser. B* 57 (1995) 289–300.
- [5] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum, NY, 1981.
- [6] B. Biswal, J. Ulmer, Blind source separation of multiple signal sources of fMRI data sets using independent component analysis, *J. Comput. Assisted Tomography* 23 (2) (1999) 265–271.
- [7] E. Bullmore, C. Long, J. Fadili, G. Calvert, F. Zalaya, T.A. Carpenter, M. Brammer, Colored noise and computational inference in neurophysiological (fMRI) time series analysis: resampling methods in time and wavelet domains, *Hum. Brain Mapp.* 12 (2001) 61–78.
- [8] S. Chen, D. Zhang, Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure, *IEEE Trans. Systems Man Cybern.* 34 (4) (2004) 1907–1916.
- [9] K. Chuang, M. Chiu, C. Lin, J. Chen, Model-free functional MRI analysis using Kohonen clustering neural network and fuzzy *c*-means, *IEEE Trans. Med. Imaging* 18 (1999) 1117–1128.
- [10] X. Descombes, F. Kruggel, D.Y. von Cramon, Spatio-temporal fMRI analysis using Markov random fields, *IEEE Trans. Med. Imaging* 17 (6) (1998) 1028–1039.
- [11] 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, *Hum. Brain Mapp.* 2 (1995) 189–210.
- [12] G.A. Hossein-Zadeh, H. Soltanian-Zadeh, B.A. Ardekani, Multiresolution fMRI activation detection using translation invariant wavelet transform and statistical analysis based on resampling, *IEEE Trans. Med. Imaging* 22 (3) (2003) 302–314.
- [13] T. Kohonen, *Self-organizing Maps*, Springer Series in Information Sciences, second extended ed., Springer, New York, 1997.
- [14] W. Lu, J.C. Rajapakse, Approach and applications of constrained ICA, *IEEE Trans. Neural Networks* 16 (1) (2005) 203–212.
- [15] W. Lu, J.C. Rajapakse, ICA with reference, *Neurocomputing* 69 (2006) 2244–2257.
- [16] M.J. McKeown, 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, *Hum. Brain Mapp.* 6 (1998) 160–188.
- [17] A. Meyer-Baese, A. Wismueller, O. Lange, Comparison of two exploratory data analysis methods for fMRI: unsupervised clustering versus independent component analysis, *IEEE Trans. Inform. Technol. Biomed.* 8 (3) (2004) 387–398.
- [18] S. Ngan, X. Hu, Analysis of functional magnetic resonance imaging data using self-organizing mapping with spatial connectivity, *Magn. Reson. Med.* 41 (1999) 939–946.
- [19] S. Ogawa, D. Tank, R. Menon, Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging, *Proc. Natl. Acad. Sci., USA* 89 (1992) 5951–5955.
- [20] D.L. Pham, J.L. Prince, Adaptive fuzzy segmentation of magnetic resonance imaging, *IEEE Trans. Med. Imaging* 18 (9) (1999) 737–752.
- [21] J.C. Rajapakse, C. DeCarli, A. McLaughlin, J.N. Giedd, A.L. Krain, S.D. Hamburger, J.L. Rapoport, Cerebral magnetic resonance image segmentation using data fusion, *J. Comput. Assisted Tomography* 20 (2) (1996) 206–218.
- [22] J.C. Rajapakse, J.N. Giedd, C. DeCarli, J.W. Snell, A. McLaughlin, Y.C. Vauss, A.L. Krain, S. Hamburger, J.L. Rapoport, A technique for single-channel MR brain tissue segmentation: application to a pediatric sample, *Magn. Reson. Imaging* 14 (9) (1996) 1053–1065.
- [23] J.C. Rajapakse, F. Kruggel, J.M. Maisog, D.Y. von Cramon, Modeling hemodynamic response for analysis of functional MRI time-series, *Hum. Brain Mapp.* 6 (1998) 283–300.
- [24] J.C. Rajapakse, J. Piyaratna, Bayesian approach to segmentation of statistical parametric maps, *IEEE Trans. Biomed. Eng.* 48 (2001) 1186–1194.
- [25] E. Salli, H.J. Aronen, S. Savolainen, A. Korvenoja, A. Visa, Contextual clustering for analysis of functional MRI data, *IEEE Trans. Med. Imaging* 20 (5) (2001) 403–414.
- [26] L. Tamm, V. Menon, C.K. Johnston, D.R. Hessel, A.L. Reiss, fMRI study of cognitive interference processing in females with fragile x syndrome, *J. Cognitive Neurosci.* 14 (2) (2002) 160–171.
- [27] A.L. Vazquez, D.C. Noll, Nonlinear aspects of the bold response in functional MRI, *Hum. Brain Mapp.* 40 (1998) 249–260.
- [28] Y. Wang, J.C. Rajapakse, Contextual modeling of functional MR images with conditional random fields, *IEEE Trans. Med. Imaging* 25 (6) (2006) 804–812.
- [29] Wellcome Department of Imaging Neuroscience, “Statistical Parametric Mapping” (<http://www.fil.ion.ucl.ac.uk/spm/>), 2004.
- [30] M.W. Woolrich, T.E.J. Behrens, C.F. Beckmann, S.M. Smith, Mixture models with adaptive spatial regularization for segmentation with an application to fMRI data, *IEEE Trans. Med. Imaging* 24 (1) (2005) 1–11.
- [31] M.W. Woolrich, M. Jenkinson, J.M. Brady, S.M. Smith, Fully Bayesian spatio-temporal modeling of fMRI data, *IEEE Trans. Med. Imaging* 24 (2) (2004) 213–231.
- [32] X. Zhong, J.C. Rajapakse, Learning functional structure from fMR images, *NeuroImage* 31 (4) (2006) 1601–1613.
- [33] J. Zhou, J.C. Rajapakse, Segmentation of subcortical brain structures using fuzzy templates, *NeuroImage* 28 (4) (2005) 927–936.
- [34] J. Zhou, J.C. Rajapakse, Extraction of fuzzy features for detecting brain activation from functional MR time-series, *Lecture Notes in Computer Science*, vol. 4243, Springer, Berlin, 2006, pp. 983–992.



Dr. Helen Juan Zhou received B.Eng. degree with First Class Honors in accelerated bachelor program in computer engineering from Nanyang Technological University, Singapore, in 2003 and the Ph.D. degree in computer engineering from the Nanyang Technological University, Singapore, in 2008. She has received undergraduate scholarship from Ministry of Education, Singapore, from October 1998 to December 2003.

Currently she is a postdoctoral Research Fellow at Bioinformatics Research Centre, Nanyang Technological University and an associate member of Asia Pacific Science and Technology Center, Sun Microsystems Inc. Her research interests include biomedical image

analysis including segmentation, activation detection, and brain connectivity inference from both anatomical and functional brain images by using machine learning techniques, dynamic Bayesian networks, fuzzy modelling, and multi-modality fusion.



Dr. Jagath C. Rajapakse is an Associate Professor at the School of Computer Engineering (SCE) and the Director of the Bioinformatics Research Centre (BIRC) at the Nanyang Technological University (NTU), Singapore. He is also a Visiting Professor at the Biological Engineering Division, Massachusetts Institute of Technology (MIT), USA. Dr. Rajapakse received B.Sc. (Eng.) degree with First Class Honors in electronic and telecommunication engineering from the University of Moratuwa (Sri Lanka), and M.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Buffalo (USA). He was a Visiting Scientist at the Max-Planck-Institute of Cognitive and Brain Sciences (Leipzig, Germany) and a Visiting Fellow at the National Institute of Mental Health (Bethesda, USA).

Dr. Rajapakse's current research interests are on gene networks, protein interactions, neural systems, and pathways. He has authored over 200 research publications in refereed journals, books, and conference proceedings in the fields of brain imaging, computational biology, and machine learning. Thompson ISI listed him among the most cited scientists of all fields. He serves as an Associate Editor of the *IEEE Transactions on Computational Biology and Bioinformatics* and of the *IEEE Engineering in Medicine and Biology Magazine*. He is presently the Chair of IAPR Technical Committee on Pattern Recognition for Bioinformatics. He was the General Chair of PRIB 2007, General Co-Chair of CIBC 2007, and Program Co-Chair of EvoBio 2007.