A neighborhood method for statistical analysis of fMRI data

by

Fayyaz Ahmad and Sung-Ho Kim

BK21 Research Report

10 - 07

August 16, 2010

DEPARTMENT OF MATHEMATICAL SCIENCES

The second stage of BK21 Fostering A World Class Talent

KAIST 한국과학기술원 Korea Advanced Institute of Science and Technology

A Neighborhood Method for Statistical Analysis of fMRI data

Fayyaz Ahmad[a], Sung-Ho Kim[*,a]

[a]*Korea Advanced Institute of Science and Technology,Department of Mathematical Sciences , 373-1, Guseong Dong, Yuseog-Gu, Daejeon 305-701, Republic of Korea.*

Abstract

In an effort to cope with the fact that functional magnetic resonance imaging (fMRI) data are spatio-temporally correlated, we propose a novel statistical method with a view to improve the detection of brain regions with increased neuronal activity in fMRI. In this method, we make use of information from neighboring voxels of a voxel, for estimation at the voxel. We examined performance of the method against the statistical parametric mapping (SPM) method using both simulated and real data. The proposed method is shown to be considerably better than the SPM in the context of Receiver Operating Characteristics (ROC) curves.

Key words: statistical parametric mapping, autoregressive model, initial values, GLM, ROI, ROC curve

1. INTRODUCTION

Functional Magnetic Resonance Imaging is based on the principles of magnetic resonance and the fact that increases in neural activity are accompanied by changes in regional cerebral blood flow (rCBF) and blood oxygenation. This blood oxygenation level dependent (BOLD) effect is the basis for most of the fMRI studies to map patterns of activation in the working human brain. The mapping and assigning of brain functions is well reflected in its

[*]Corresponding author. Tel +82-42-350-2737, Fax +82-42-350-2710

Email addresses: fayyaz@kaist.ac.kr (Fayyaz Ahmad), sung-ho.kim@kaist.edu (Sung-Ho Kim)

URL: http://mathsci.kaist.ac.kr/∼slki (Sung-Ho Kim)

name, functional MRI.

fMRI has been in use for investigator of a variety of neuronal processes from activities in the primary visual and auditory cortices to cognitive functions such as sensation or listening. In an fMRI experiment, baseline images are scanned of equal spaced time points while the subject is at rest and activation images are acquired while the subject is performing a particular task. Lueck et al.[1989] and Friston et al.[1994] are among the first to develop a statistical parametric mapping package(SPM: www.fil.ion.ucl.ac.uk/spm) based on a voxel-wise analysis.

There are two major problems related to SPM: 1) modeling evoked hemodynamic response function (HRF) in fMRI time series data; and 2) autocorrelation of random errors, (i.e., successive fMRI scans are not independent). A linear time-invariant or a convolution model was proposed to overcome the first problem [Kiebel et al., 2003]. It is difficult to find the correct form of the convolution kernel to accurately describe the hemodynamic response function. The second problem is important in analysis of fMRI scans where successive time series scans are not independent. A solution to the problem was given by Worsely and Friston(2002) and provided a general framework for the estimation of serial correlations among error terms.

In the SPM, we estimate the error covariance matrix, $\sigma^2 V$ say, by employing the restricted maximum likelihood method [Harville, 1974]. In the method, constraints are imposed on V in the form of $V = \sum_{l=1}^{2} \lambda_l Q_l$ with $Q_1 = I$, $Q_2 = \left(e^{-|i-j|} : i \neq j\right)$, where I is an identity matrix and the hyperparameters (error variance components) λ_l are estimated based on pooled data from all the interested voxels. The V is used to de-correlate the data of all the voxels of the brain and the σ is estimated for each voxel.

The fMRI noise is autocorrelated and functional imaging data have some spatial correlation. This correlation is further enhanced by some operations such as smoothing and reslicing fMRI data, and also, fMRI data of low resolution from an individual voxel will contain some signal from the tissue around that voxel. This spatio-temporal autocorrelation [Valdes-Sosa, 2004] has led us to propose a method in which we incorporate estimation results at neighboring voxels of a voxel, v say, in the estimation at voxel v. This proposed method is shown to increase substantially the statistical significance

of activated regions as compared to the traditional SPM approach.

2. THE PROPOSED METHOD

Many methods are available for the statistical analysis of fMRI data that range from a simple linear model for the response and a global first-order autoregressive model [Watson, 1955; Seber, 1977 and Worsely et al., 2002] for the temporal errors, to a more sophisticated non-linear model for the response with a local state space model for the temporal errors [Purdon et al., 2001]. In the analysis of fMRI data, the SPM adopts an AR(1) model for the noise in the original data, whose autocorrelation coefficient, ρ, is assumed to be constant across the voxels of the brain rather than varying spatially as proposed in our method.

There is a considerable evidence that the error covariance matrix is not homogeneous [Purdon et al., 2001]. In the SPM, we analyze data with regard to temporal correlation but not consider any relationship between voxels to decorrelate the fMRI data. The estimates of ρ and the variance of the white noise (σ_ξ^2) of all the voxels of a slice are shown in Fig. 9 and 10 which are obtained from real data. Inspired by this variation of the estimates we propose an efficient, pre-whitening strategy [Worsley et al., 2002] with estimation made at every voxel in the brain. We will call this method a neighborhood method (NH). It also incorporates the spatial correlation between neighboring voxels and improves the detection accuracy of active voxels. The NH increases substantially the statistical significance of activated regions, which makes it possible to decide with higher confidence if a certain brain region is activated or not. We compare our approach to functional imaging with the SPM method [Friston et al., 1995a; Worsley et al., 1995].

The NH algorithm is as follows:

1. The general linear model which we apply at voxel v is given by

$$Y_v = X_v\beta_v + \varepsilon_v, \quad \varepsilon_v \sim N(0, V_v), \tag{1}$$

   where X_v is a design matrix.
2. Suppose the parameter estimation begins at voxel v_1. We obtain $\hat{\varepsilon}_{v_1} = Y_{v_1} - \hat{Y}_{v_1}$ where $\hat{Y}_{v_1} = X_{v_1}\hat{\beta}_{v_1}$ and $\hat{\beta}_{v_1} = (X_{v_1}^T X_{v_1})^{-1} X_{v_1}^T Y_{v_1}$. We calculate $\hat{V}_{v_1}$ by adopting the first order autoregressive model proposed by Bullmore et al.[1996]. In this model images are equally spaced in time

and the errors from previous images, $\varepsilon_{v_1}(t-1)$, are mixed up with a white noise $\xi_{v_1}(t)$ into the error of the current image, $\varepsilon_{v_1}(t)$:

$$\varepsilon_{v_1}(t) = \rho\varepsilon_{v_1}(t-1) + \xi_{v_1}(t) \tag{2}$$

where $|\rho| < 1$ and $\{\xi_{v_1}(t)\}$ forms a sequence of independent and identically distributed errors with $\xi_{v_1}(t) \sim N(0, \sigma^2_{\xi_{v_1}})$. The correlation matrix of the first order autoregressive model is given by

$$V_{v_1} = \rho_{v_1}^{|s|} \frac{\sigma^2_{\xi_{v_1}}}{1-\rho^2_{v_1}}, \tag{3}$$

where $\rho^{|s|}$ is a square matrix whose (i,j)th entry is $\rho^{|i-j|}$, $i, j = 1, \cdots, s$. i.e.,

$$\rho^{|s|} = \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{s-1} \\ \rho & 1 & \rho & \dots & \rho^{s-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho^{s-1} & \rho^{s-2} & \rho^{s-3} & \dots & 1 \end{pmatrix}$$

3. Calculate the de-correlating matrix $\hat{W}_{v_1}$ as $\hat{W}_{v_1} = \hat{V}_{v_1}^{-1/2}$. Multiply the general linear model (1) by $\hat{W}_{v_1}$:

$$\hat{W}_{v_1} Y_{v_1} = \hat{W}_{v_1} X_{v_1} \beta_{v_1} + \hat{W}_{v_1} \varepsilon_{v_1}$$

or

$$Y^*_{v_1} = X^*_{v_1} \beta^*_{v_1} + \varepsilon^*_{v_1}, \quad \varepsilon^*_{v_1} \sim N(0, \sigma^2 I), \tag{4}$$

where the terms matching in the above two equations are the same, e.g., $\hat{W}_{v_1} Y_{v_1} = Y^*_{v_1}$. $\beta^*_{v_1}$ is then estimated as

$$\hat{\beta}^*_{v_1} = (X^{*T}_{v_1} X^*_{v_1})^{-1} X^{*T}_{v_1} Y^*_{v_1} \tag{5}$$

4a. For a neighboring voxel v_2 of v_1, set initial values of the estimates as:

$$\begin{aligned} \hat{\beta}^{*(0)}_{v_2} &= \hat{\beta}^*_{v_1}, \\ \hat{\varepsilon}^{*(0)}_{v_2} &= \hat{W}_{v_1} Y_{v_2} - \hat{W}_{v_1} X_{v_2} \hat{\beta}^{*(0)}_{v_2}, \\ \hat{V}^{*(0)}_{v_2} &= \hat{\rho}_{v_1}^{|s|} \frac{\hat{\sigma}^{2(0)}_{\xi_{v_2}}}{1-\hat{\rho}^2_{v_1}}, \\ \hat{W}^{(0)}_{v_2} &= (\hat{V}^{*(0)}_{v_2})^{-1/2}. \end{aligned}$$

where $\hat{\sigma}^{2(0)}_{\xi_{v_2}}$ is computed based on $\hat{\varepsilon}^{*(0)}_{v_2}$ by using equation (2).

4b. Repeat updating $\hat{\beta}^{*(k-1)}_{v_2}$, $\hat{\varepsilon}^{*(k-1)}_{v_2}$, and $\hat{W}^{(k-1)}_{v_2}$ until convergence takes place (e.g., $|\hat{\beta}^{*(k)}_{v_2} - \hat{\beta}^{*(k-1)}_{v_2}| \leq 0.001$, $|\hat{\varepsilon}^{*(k)}_{v_2} - \hat{\varepsilon}^{*(k-1)}_{v_2}| \leq 0.001$ and $||\hat{W}^{(k)}_{v_2,ij} - \hat{W}^{(k-1)}_{v_2,ij}|| \leq 0.001$):

$$\begin{aligned}
X^*_{v_2} &= \hat{W}^{(k-1)}_{v_2} X_{v_2} \\
Y^*_{v_2} &= \hat{W}^{(k-1)}_{v_2} Y_{v_2} \\
\hat{\beta}^{*(k)}_{v_2} &= (X^{*T}_{v_2} X^*_{v_2})^{-1} X^{*T}_{v_2} Y^*_{v_2}, \\
\hat{\varepsilon}^{*(k)}_{v_2} &= Y^*_{v_2} - X^*_{v_2} \hat{\beta}^{*(k)}_{v_2}, \\
\hat{V}^{*(k)}_{v_2} &= \hat{\rho}^{|s|}_{v_2} \frac{\hat{\sigma}^{2(k)}_{\xi_{v_2}}}{1 - \hat{\rho}^2_{v_2}}, \\
\hat{W}^{(k)}_{v_2} &= (\hat{V}^{*(k)}_{v_2})^{-1/2},
\end{aligned}$$

4c. If a voxel v has more than one voxels where estimation is already made, then we take averages of the estimates from the neighboring voxels for the initial values of the parameters β, ε, V, and W. In other words, the right-hand sides of the equations in step 4a are replaced with the corresponding averages of the estimates.

5. Apply step 4 to all the voxels of the brain that are involved in the data.

3. EXPERIMENTAL RESULTS

3.1. Using Simulated fMRI Data

In order to evaluate our method, we generated a Gaussian artificial activation in 4-D fMRI, with the size of $6 \times 5 \times 5$ voxels per volume in four locations and 80 volumes. The 80 volumes of fMRI had alternating blocks of 40 non-active and 40 active volumes, beginning with non-active volumes. The activations were made in four different regions of interest (ROI) as shown in Table 1 and in the three slices, 17, 20, and 4, as shown in Fig. 1. The means and standard deviations (SD) of voxel values of 40 active volumes corresponding to the four ROI's are also shown in Table 1. The artificial activation signals, Z_{it} as shown in Fig. 2, in four ROI's are generated from Gaussian distributions according to the following model:

$$Y_{it} = X_t\beta_i + \varepsilon_{it}, \tag{6}$$
$$Z_{it} = Y_{it} + a_{it}\ ;\ a_{it} \sim N(\mu, 25), \quad i = 1, \cdots, 150 \tag{7}$$

The 4-D fMRI data contain BOLD/EPI acquisitions consisting of 35 slices (image volume size in voxels: $x = 128, y = 128, z = 35$). 80 acquisitions or volumes were made: 8 blocks and 10 reps in each block for active and non-active conditions. In the 10 repetitions block, we assume a task stimulus, and rest conditions for the next 10 repetitions block. Successive blocks alternated between task and rest and the design matrix settled up a blocked paradigm for the simulation data.

Table 1: Artificial activation in four different ROI's

ROI	Voxel	Positions			Mean	SD	Artificial activation
		X	Y	Z			
1	150	51-56	65-69	14-18	729.66	99.51	$N(20, 25)$
2	150	74-79	65-69	14-18	733.54	105.23	$N(20, 25)$
3	150	63-69	95-99	18-22	984.44	183.90	$N(15, 25)$
4	150	63-69	31-35	4-8	844.10	130.07	$N(15, 25)$

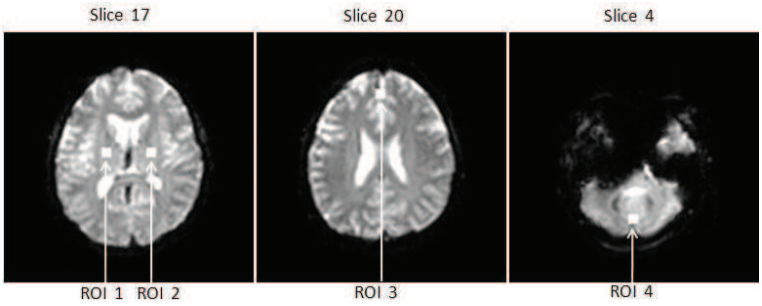


Figure 1: Three slice views of artificial activation areas of four ROI's in the brain.

A volume of SPM {t} map was constructed with a task-and-rest condition of design matrix corresponding to $p < 0.01$ (corrected) level by using the SPM method as shown in Fig. 3 and by using the NH method as shown

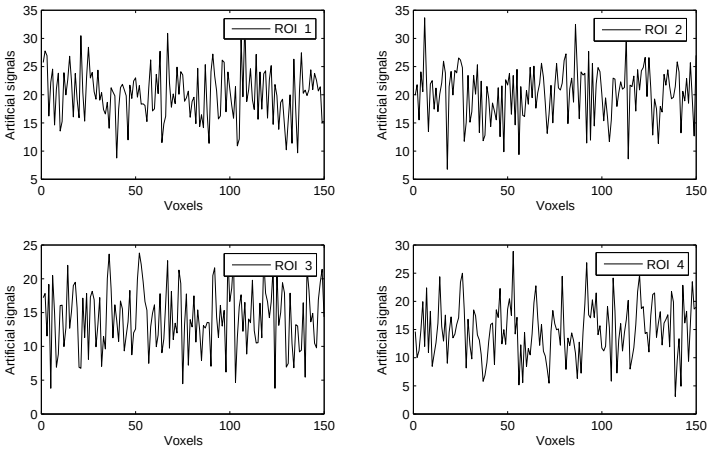
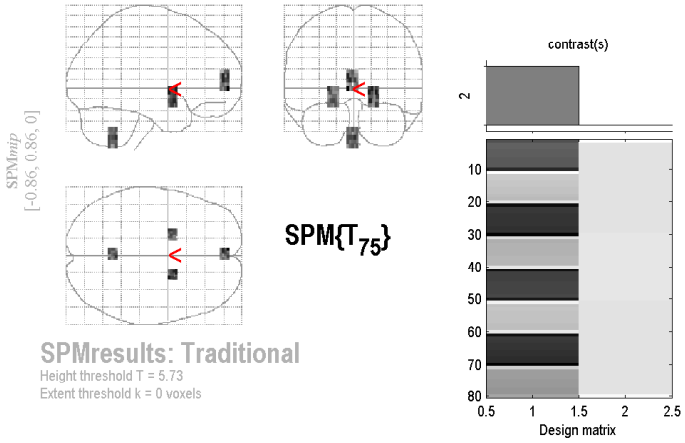


Figure 2: Artificial signals at four ROI's

in Fig. 4. In the figures the images to the left show a so-called "glass brain" view of activated voxels of the whole brain. The bar-graph on the right represents the design matrix of task and rest conditions and above it is a graphical representation of the contrast $\mathbf{c}$=[1 0]. The tables in the figures show statistical 't' maps of activated voxels corresponding to $p < 0.01$ (corrected). The number of voxels in the activated regions clearly shows that the NH method increases substantially the statistical significance of the four activated regions.

We will make use of the well-known Receiver Operating Characteristics (ROC) curve analysis [Kim et al., 2005] to compare the SPM and NH methods. A ROC curve is a graphical representation of the true positive rate (sensitivity) versus the false positive rate (1 - specificity) for a binary classifier system over a range of its discrimination threshold. The frequencies of true positive and false positive voxels are obtained by applying four different threshold levels to the SPM {t} maps of simulation data for both methods. The sensitivity or cumulative rates of true and false positive voxels corresponding to these threshold levels for both methods are estimated as shown in Table 2 and the ROC curves are fitted to four of the bivariate pairs, as shown in Fig. 5(a). The area under the ROC curve from each of the two methods is 0.9316 and 0.9218 respectively for the NH and the SPM method. The difference of the area between the two methods is relatively small since



Statistics: *p-values adjusted for search volume*

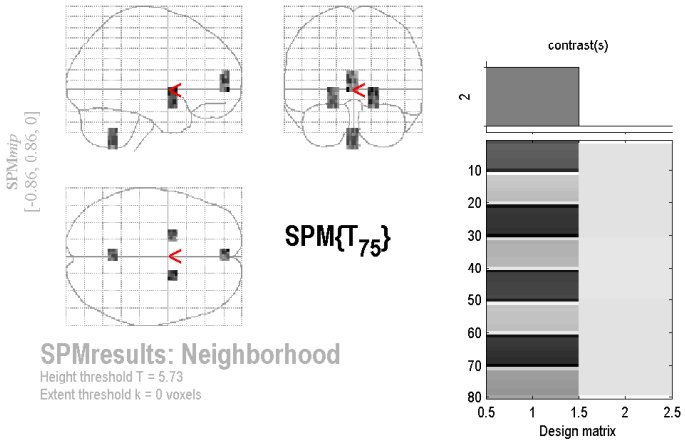
set-level		cluster-level			voxel-level					x,y,z {mm}		
p	c	$p_{corrected}$	k_E	$p_{uncorrected}$	$p_{FWE-corr}$	$p_{FDR-corr}$	T	(Z_e)	$p_{uncorrected}$			
0.000	**7**	**0.000**	**69**	**0.000**	**0.000**	**0.000**	**13.12**	**Inf**	**0.000**	**-6**	**59**	**0**
					0.000	0.000	11.53	Inf	0.000	-6	54	12
					0.000	0.000	10.48	Inf	0.000	1	59	8
		0.000	**103**	**0.000**	**0.000**	**0.000**	**11.93**	**Inf**	**0.000**	**22**	**1**	**-4**
					0.000	0.000	11.74	Inf	0.000	15	8	-4
					0.000	0.000	11.56	Inf	0.000	16	6	-12
		0.000	**66**	**0.000**	**0.000**	**0.000**	**10.98**	**Inf**	**0.000**	**1**	**-51**	**-48**
					0.000	0.000	9.86	7.84	0.000	-6	-58	-52
					0.000	0.000	9.83	7.83	0.000	-1	-51	-56
		0.000	**73**	**0.000**	**0.000**	**0.000**	**10.89**	**Inf**	**0.000**	**-25**	**1**	**0**
					0.000	0.000	10.08	Inf	0.000	-20	3	-12
					0.000	0.000	8.07	6.82	0.000	-16	6	0
		0.000	**1**	**0.002**	**0.000**	**0.000**	**9.64**	**7.74**	**0.000**	**3**	**-58**	**-56**
		0.000	**1**	**0.002**	**0.000**	**0.000**	**6.56**	**5.81**	**0.000**	**-25**	**1**	**-16**
		0.000	**1**	**0.002**	**0.001**	**0.000**	**6.43**	**5.72**	**0.000**	**23**	**8**	**0**

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 5.73, p = 0.000 (0.010)
Extent threshold: k = 0 voxels, p = 1.000 (0.010)
Expected voxels per cluster, <k> = 0.084
Expected number of clusters, <c> = 0.11
Expected false discovery rate, <= 0.00

Degrees of freedom = [1.0, 75.0]
Smoothness FWHM = 3.1 3.0 5.8 {mm} = 1.8 1.7 1.5 {voxels}
Search vol: 1195466 cmm; 101023 voxels; 19229.9 resels
Voxel size: [1.7, 1.7, 4.0] mm (1 resel = 4.58 voxels)

Figure 3: Results of the test of artificial activations by using the SPM method. (Top right) This is an image representation of the design matrix. The contrast is displayed above the column of design matrix that corresponds to the activation effect $\mathbf{c}$=[1 0]. (Top left) SPM {t}: This is a maximum intensity of the projection of SPM {t}. The display format is standard and provides three views of the brain from the front, below, and right-hand side. The grayscale squared and rectangular shapes are the activated areas of our ROIs that are described in the atlas of Talairach. (Bottom) Tabular data are presented for "significant" regions $p < 0.01$ (corrected). The location of the maximal voxel in each region (right column) with the size of region (K_E) or cluster up to three t and z maxima. For each maximum the significance is assessed in term of p–values, t–values and z–values. In this figure there are mainly four (ROI) significant regions with $p < 0.01$ (corrected).



Statistics: *p-values adjusted for search volume*

set-level		cluster-level			voxel-level					x,y,z {mm}		
p	*c*	$p_{corrected}$	k_E	$p_{uncorrected}$	$p_{FWE-corr}$	$p_{FDR-corr}$	*T*	(Z_e)	$p_{uncorrected}$			
0.000	**4**	**0.000**	**73**	**0.000**	**0.000**	**0.000**	**14.48**	**Inf**	**0.000**	**-6**	**59**	**0**
					0.000	0.000	12.20	Inf	0.000	-4	54	8
					0.000	0.000	10.12	Inf	0.000	-1	59	16
		0.000	**112**	**0.000**	**0.000**	**0.000**	**13.55**	**Inf**	**0.000**	**18**	**8**	**0**
					0.000	0.000	12.98	Inf	0.000	22	1	-4
					0.000	0.000	12.76	Inf	0.000	16	6	-12
		0.000	**89**	**0.000**	**0.000**	**0.000**	**11.86**	**Inf**	**0.000**	**-25**	**6**	**0**
					0.000	0.000	10.83	Inf	0.000	-18	3	-16
					0.000	0.000	7.81	6.66	0.000	-16	6	0
		0.000	**82**	**0.000**	**0.000**	**0.000**	**11.17**	**Inf**	**0.000**	**-1**	**-54**	**-52**
					0.000	0.000	11.04	Inf	0.000	3	-51	-40
					0.000	0.000	9.01	7.39	0.000	-6	-52	-40

table shows 3 local maxima more than 8.0mm apart

Height threshold: T = 5.73, p = 0.000 (0.010)
Extent threshold: k = 0 voxels, p = 1.000 (0.010)
Expected voxels per cluster, <k> = 0.081
Expected number of clusters, <c> = 0.11
Expected false discovery rate, <= 0.00

Degrees of freedom = [1.0, 75.0]
Smoothness FWHM = 3.1 3.0 5.8 {mm} = 1.8 1.7 1.4 {voxels}
Search vol: 1195466 cmm; 101023 voxels; 19885.9 resels
Voxel size: [1.7, 1.7, 4.0] mm (1 resel = 4.43 voxels)

Figure 4: Results of the test of artificial activations by using the NH method. The parts of this figure are as explained in Fig. 3.

the artificial signals are added to only 600 voxels (150 voxels for each ROI). The ROC curve and Table 2 show that the NH method has: (1) larger true positive rate; (2) lower false positive rate; (3) larger frequencies of true positive; (4) a slightly larger area under the ROC curve by the NH than by the SPM.

In the NH method, we assume that each voxel is highly correlated with the neighboring voxels partly due to the smoothing of the data images and we make use of the information from neighboring voxels of a voxel, for estimation at the voxel. The distinction of the ROC curves between the NH and the SPM methods becomes more apparent for real data as we will see in the next subsection.

Table 2: Sensitivity analysis with the SPM and NH methods

	Observed Frequencies				Cumulative Rates			
Threshold level	False Positive		True Positive		False Positive		True Positive	
SPM{t}	SPM	NH	SPM	NH	SPM	NH	SPM	NH
$\geq$5.75	1	1	311	351	0.0667	0.0417	0.7814	0.7817
$\geq$5.50	2	3	44	50	0.2000	0.1667	0.8920	0.8931
$\geq$5.25	5	7	23	27	0.5333	0.4583	0.9497	0.9532
$\geq$5.0	7	13	20	21	1	1	1	1

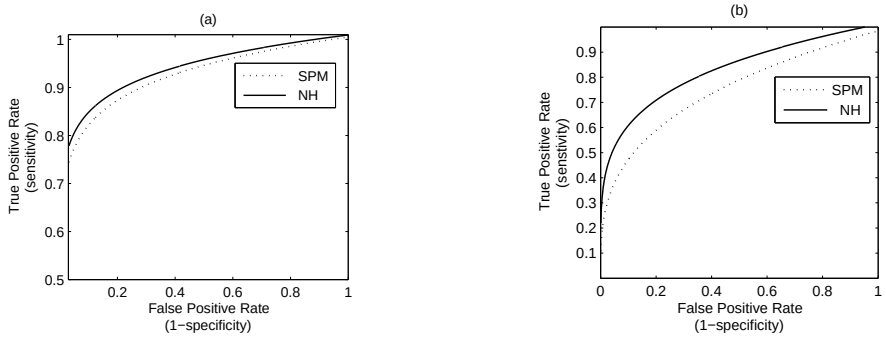


Figure 5: ROC curves by the SPM and NH methods. Panel (a) shows the result from simulation data and the result from real data is in panel (b).

3.2. Using Real fMRI Data

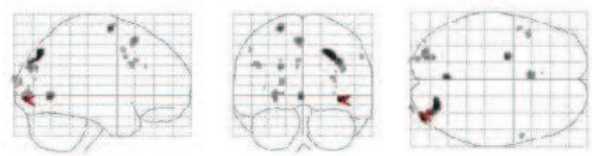
We also compared the SPM and the NH methods with real fMRI data obtained from a visual block design experiment, which is available in the public domain at http://cnl.web.arizona.edu/spm.htm. The three condition ("study") block images data were acquired on a GE 1.5T Sigma 5x Whole-body Echospeed Horizon System. The whole brain BOLD/EPI acquisition consisted of 17 slices, each 5 mm thick, with a 1 mm skip (image volume size in voxels: $x = 64, y = 64, z = 17$; voxel size: 3.44 mm $\times$ 3.44 mm $\times$ 6 mm; FOV=220). The acquisition took 160 seconds, with the scan-to-scan repetition time (TR) set to 2 secconds. A total of 80 acquisitions were made: 8 blocks of 20 seconds each (i.e., 8 blocks of 10 reps each). During each 20 second block, we presented 4 stimuli of bird pictures, each for 5 seconds. Successive blocks alternated between hard and easy learned birds and a control condition of familiar birds (crows and swans), starting with learned birds. The pattern was as in Table 3.

Table 3: Block design of fMRI "study" data

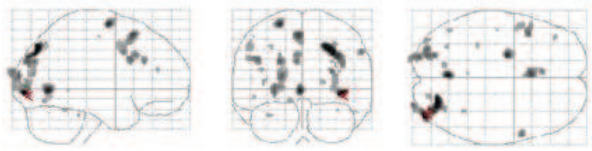
Block 1	Block 2	Block 2	Block 4	Block 5	Block 6	Block 7	Block 8
20 sec	20 sec	20 sec	20 sec	20 sec	20 sec	20 sec	20 sec
10 TRs	10 TRs	10 TRs	10 TRs	10 TRs	10 TRs	10 TRs	10 TRs
4 easy	4 controls	4 hard	4 controls	4 easy	4 controls	4 hard	4 controls

A "2D" T1 anatomical (same plane and section as the BOLD image; Series 2) and a "3D" image (124 Sagittal slices, Series 4) were also acquired to be used as structural images. We obtained SPMs of the real fMRI data after pre-processing with smoothing {FWHM = 7 mm} using the NH and the SPM methods with $p < 0.05$ (corrected) and contrast $\mathbf{c}$=[1 -2 1]. Results are displayed in the "glass-brain" views. Gray areas reflect the easy and hard conditions causing higher brain activities than the control condition as shown in Fig. 6. An increase in the cerebral blood flow is detected in the temporal lobe, the thalamus, and the Wernicke area as a result of "study" conditions. Small scattered areas of activity are seen in the projection of the anterior commissure, sensory cortex, and frontal and temporal lobes.

Fig. 7 depicts the SPMs overlaid on the rendering of the structural MRI. The brain images in panel (a) depicts SPMs on the rendering obtained by the SPM method and the images in panel (b) by the NH.



(a)



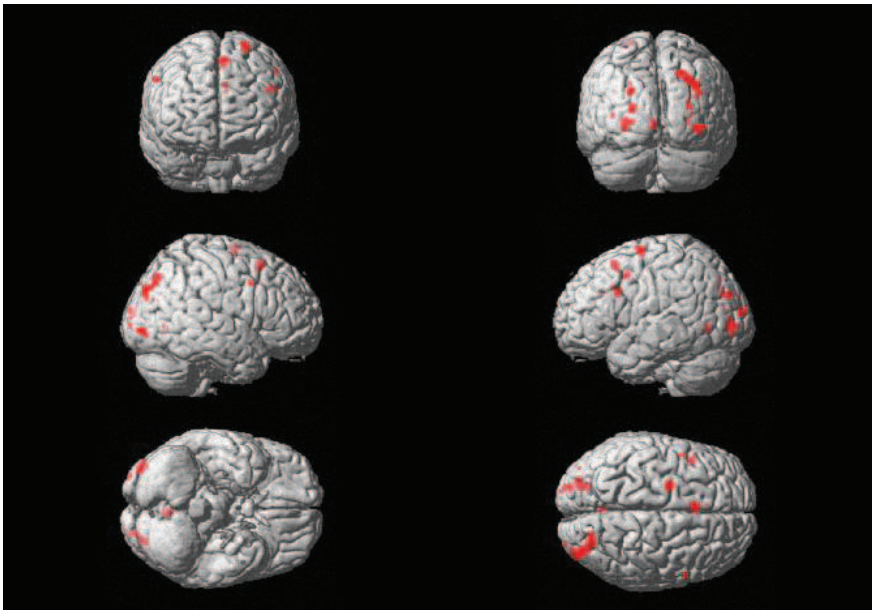
(b)

Figure 6: Three plane fMRI "glass-brain" images. The detection of the active regions using the traditional SPM approach (a), and the NH method (b). Gray areas reflect the higher activities in the brain as a result of the easy and hard conditions than the control condition.

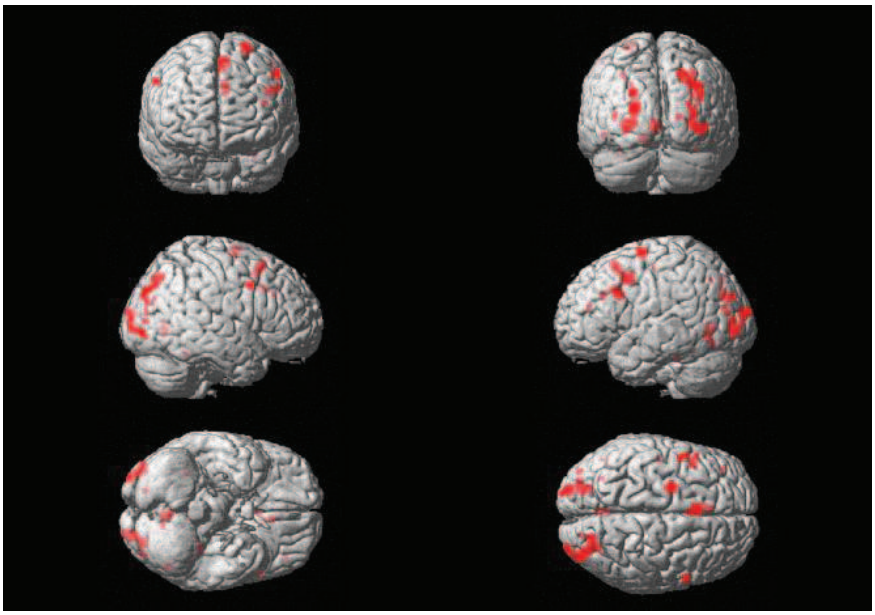
A comparison of the two methods is also made with $p < 0.05$ (corrected). The SPM method detected 8814 voxels whereas the NH method detected 12710 voxels; the p-values of all the voxels are shown in Fig. 8. The figure shows that the NH method has smaller p-values of activated voxels than the SPM. The ROC curves in Fig. 5(b) shows that the NH method is superior to the SPM method and the areas under the ROC curves by the NH and the SPM are 0.8238 and 0.7356, respectively. The p-values by the NH method are dispersed over a much wider range than the SPM method which is reflected in the ROC curves with larger true positive rates and smaller false positive rates for the NH method.

4. DISCUSSION

In the SPM, the error covariance matrix, $\sigma^2 V$, involves hyperparameters which do not vary over voxels, but the σ is estimated at each voxel to best match the error covariance at that voxel. On the other hand, we admit more heterogeneity of the covariance matrix in the NH method.



(a)



(b)

Figure 7: Regions are rendered in red on the MNI template of SPM of activated regions of the brain as a result of "higher active" of easy and hard conditions than the control condition. Active regions detected with the use of the SPM method (a), and the NH method (b) correspond to $p < 0.05$ (corrected).

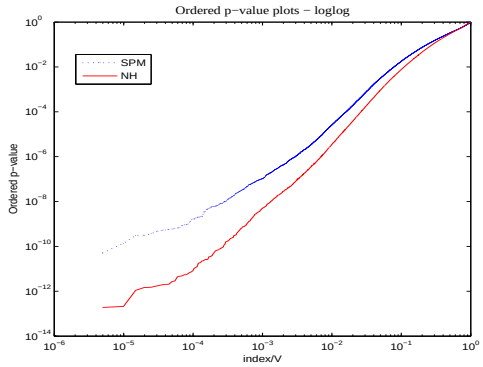


Figure 8: The p-values of all the voxels corresponding to $p(< 0.05)$ corrected FDR by using the SPM method (blue line) and the NH method (red line)

Fig. 9 and 10 show the estimates of ρ and σ_ξ^2 in (3) of all the voxels of slice $Z = 35$, which are estimated based on the real fMRI data. The fluctuation of the values of the estimates across the voxels called our attention to a method which can incorporate the fluctuation into the estimation procedure of the parameters in the model (1). Under the assumption that the estimates may change smoothly across neighboring voxels, we used the estimates of neighboring voxels of a voxel for the initial values of the estimates for the voxel. The variation of the parameter values ρ and σ_ξ^2 across voxels and the assumed smoothness among neighboring voxels are reflected in the estimation procedure of the NH method.

Worsley et al. [2002] used pre-whitening strategy with a spatially varying V which is repeated at every voxel to estimate β in (1) without using information from neighboring voxels. Zhang et al. [2006] used the matrix $V^{-1/2}$ to estimate parameter β of whitened model and then applied Durbin-Watson (AR(1) correlation test) on the residuals of whitened model to improve on the accuracy of the autocorrelation model. The variation of the autocorrelation coefficient (ρ) calls for the need for autocorrelation modeling with initial estimates borrowed from neighboring voxels in order to attain more accurate inferences at every voxel.

In the analysis of simulated data, the NH method detects, over four ROIs, 8%, 20%, 6% and 22% more activated voxels as shown in Table 4 than the

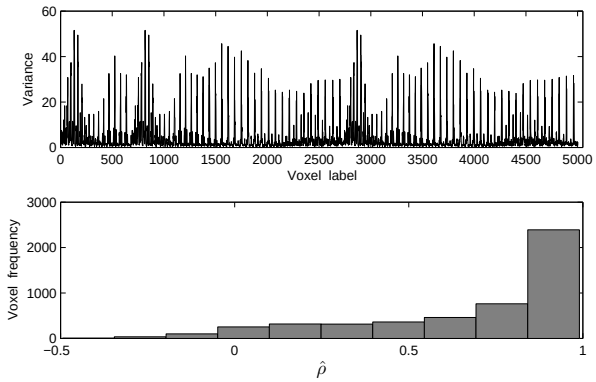


Figure 9: $\hat{\sigma}^2_\xi$ and $\hat{\rho}$ (histogram) values of the voxels in the slice, $Z = 35$ of real data. The x-axis is of the $\hat{\rho}$ values for the second graph and is of the voxel labels for the first graph. The estimates are by the NH method.

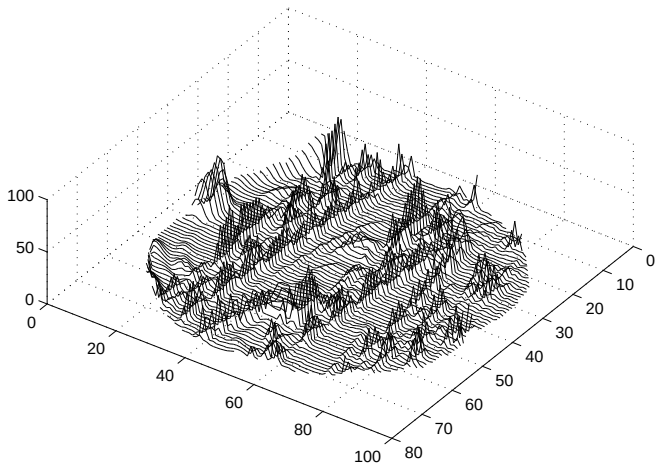


Figure 10: A bird-view of the $\hat{\sigma}^2_\xi$ values of all the voxels in the slice, $Z = 35$ of real data.

Table 4: Detection of activated voxels in the four ROI's

ROI	Activated voxels		Percentage of more detection
	SPM	NH	
1	104	112	8 %
2	74	89	20 %
3	69	73	6 %
4	67	82	22 %
Total	314	356	14 %

SPM method, and with smoothing {FWHM = 7mm} of real data detects 40% more activated voxels. The proposed method detects on average 14% more of the activated voxels of the simulated data than the SPM with a better ROC performance. The true positive rate of activated voxels of simulated data without smoothing showed validity of the proposed method and it is more apparent as far as the real data are concerned. The ROC curves and the p-values of activated voxels show that the NH method is superior than the SPM method whether they are based on the simulated data or the real data.

5. CONCLUSION

In this paper, we proposed a method in an effort to cope with the heterogeneity of the parameters, in particular the variance-covariance structure of the noise in the fMRI data. Under the assumption that the parameter estimates do not change abruptly between neighboring voxels, we employed an estimate-transfusion approach between neighboring voxels by using the estimates from neighboring voxels as initial values of the estimates for their new neighboring voxels. Since the intial values may affect the final result of the estimate [Wu, 1983 and Kim, 2002], it is desirable that we apply the estimate-transfusion approach to obtain the estimates that are affected by the estimates from neighboring voxels.

In both of the experiments, one with semi-artificial data and the other with real data, we showed superiority of the proposed NH method over the

traditional SPM method in the context of the ROC curve. In the traditional SPM, we assume a variance-covariance structure contains hyperparameters which do not vary over voxels. This may possibly deteriorate the detection accuracy of the activated voxels where the noises are relatively small. This kind of undesirable phenomenon can be avoided by applying the NH method.

The activation indexes of voxels that are obtained by the NH can be used for Bayesian inferences on the voxel activity by imposing a prior distribution on the voxel activity for a set of voxels [Everitt et al. 1999; Hartving et al. 2000; Woolrich et al. 2005].

References

Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS (1993): Processing strategies for time-course data sets in functional MRI of the human brain. Magn. Reson. Med. 30: 161-173.

Bullmore ET, Rabe-Hesketh S, Morris RG, Steven CR, Gregory L, Gray JA, Brammer MJ (1996): Functional magnetic resonance image analysis of a large-scale neurocognitive network. NeuroImage 4:16-33.

Birn RM, Cox RW, Bandettini PA (2004): Functional MRI experimental designs and processing strategies for studying brain activation associated with overt responses. NeuroImage 23:1046-1058.

Boynton GM, Engel SA, Glover GH, Heeger DJ (1996): Linear systems analysis of functional magnetic resonance imaging in human. J Neurosci 16:4207-4221.

Draper NR, Smith H (1981): Applied regression analysis(3rd edition). Wiley Academic Press, New York.

Everitt B, Bullmore E (1999): Mixture model mapping of brain activation in functional magnetic resonance images. Human Brain Mapping 7:1-14.

Eliassen JC, Souza T, Sanes JN (2003): Experience-dependent activation patterns in human brain during visual-motor associative learning. Journal of Neuroscience 23:10540-10547.

Friston KJ, Glaser DE, Henson RNA, Kiebel S, Phillips C, Ashburner J (2002): Classical and bayesian inference in neuroimaging: applications. NeuroImage 16:484-512.

Friston KJ (2002): Bayesian estimation of dynamical systems: an application to fMRI. NeuroImage 16: 513-530.

Frackowiak, Friston KJ, Frith C, Dolan R, Price CJ, Zeki Z, Ashburner J, Penny WD (2003): Human brain function (2nd edition). Academic Press, Oxford, UK.

Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SC, Frackowiak RS, Turner R (1995): Analysis of fMRI time series revisited. NeuroImage 2:45-53.

Friston KJ, Jezzard PJ, Turner R (1994): Analysis of functional MRI time–series. Human brain mapping 1:153-171.

Glover GH (1999): Deconvolution of impulse response in event-related BOLD fMRI. Neuroimage 9:416-429.

Genovese CR, Lazar NA, Nichols T (2002): Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15:870-878.

Harville DA (1974): Bayesian inference for variance components using only error contrasts. Biometrika 61:383-385.

Hartving N, Jensen J (2000): Spatial mixture modelling of fMRI data. Human Brain Mapping 11(4):233-248.

Harrison L, Penny WD, Friston KJ (2003): Multivariate autoregressive modeling of fMRI time series. NeuroImage 19:1477-1491.

Holmes A, Poline JB, Friston KJ (1997): Characterising brain images with the general linear model. In: Frackowiak , Friston, KJ, Frith CD, Dolan RJ, Mazziotta JC (2nd edition.) Human brain function, pp59-84. Academic Press, San Diego.

Kim HY, Hae Yong, Giacomantone JO (2005): A new technique to obtain clear statistical parametric map by applying anisotropic diffusion to fMRI. ICIP05 3:724-727.

Kim SH (2002): Calibrated intials for an EM applied to recursive models of categorical variables. Copmutational Statistical and Data Analysis 40(1):97-110.

Lueck CJ, Zeki S, Friston KJ (1989): The colour centre in the cerebral cortex of man. Nature 340:386-389.

McCabe K, Houser D, Ryan L, Smith V, Trouard T (2001): A functional imaging study of cooperation in two–person reciprocal exchange. PNAS 98:11832-11835.

Neil C, Schwertman (1978): A note on the geisser-greenhouse correction for incomplete data split-plot analysis. Journal of the American Statistical Association 73:393-396.

Purdon PL, Solo V, Weissko RM, Brown E (2001): Locally regularized spatiotem poral modeling and model comparison for functional MRI. NeuroImage 14:912-923.

Richard SJ, Frackowiak, Friston KJ, Raymond JD, Cathy JP, William P (2003): The general linear model. In: Human brain function (2nd edition). Academic Press pp725-760, Oxford, UK.

Satterthwaite FE (1946): An approximate distribution of estimates of variance components. Biometrics 2:110-114.

Seber GAF (1977): Linear regression analysis. Wiley Press, New York, USA.

Valdes-Sosa PA (2004): Spatio-temporal autoregressive models defined over brain manifolds. Neuroinformatics 2:239-250.

Watson GS (1955): Serial correlation in regression analysis. Biometrika 42:327-341.

Woolrich MW, Behrens TEJ, Beckmann CF, Smith SM (2005): Mixture models withadaptive spatial regularization for segmentation with an application to fMRI data. IEEE Trans Med Imaging 24(1):1-11.

Worsley KJ (2002): Non-stationary FWHM and its effect on statistical inference for fMRI data. NeuroImage 16:779780.

Worsley KJ, Friston KJ (1995): Analysis of fMRI time-series revisited - again. NeuroImage 2:173-181.

Worsley KJ, Liao C, Aston J, Petre V, Duncan GH, Morales F, Evans AC (2002): A general statistical analysis for fMRI data. NeuroImage 15:1-15.

Wu CFJ (1983) On the converengence properties of the EM algorithm. The Annals of Statistics 11(1):95-103.

Zhang H, Luo WL, Nichols TE (2006): Diagnosis of single-subject and group fMRI data with SPMd. Human brain mapping 27:442-451.