ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 34 (2007) 550–564

Spatial accuracy of fMRI activation influenced by volume- and surface-based spatial smoothing techniques

Hang Joon Jo,[a] Jong-Min Lee,[a,*] Jae-Hun Kim,[a] Yong-Wook Shin,[a] In-Young Kim,[a] Jun Soo Kwon,[b] and Sun I. Kim[a]

[a]*Department of Biomedical Engineering, Hanyang University, Sungdong P.O. Box 55, Seoul 133-605, South Korea*
[b]*Department of Psychiatry, Seoul National University College of Medicine, South Korea*

Received 15 May 2006; revised 6 September 2006; accepted 29 September 2006
Available online 15 November 2006

As improvements in cortical surface modeling allowed accurate cortical topology in brain imaging studies, surface-based methods for the analysis of functional magnetic resonance imaging (fMRI) were introduced to overcome the topological deficiency of commonly used volume-based methods. The difference between the two methods is mainly due to the smoothing techniques applied. For practical applications, the surface-based methods need to quantitatively validate the accuracy of localizing activation. In this study, we evaluated the spatial accuracy of activation detected by the volume- and surface-based methods using simulated blood oxygenation level-dependent (BOLD) signals and MRI phantoms focusing on the influence of their smoothing techniques. T1- and T2-weighted phantoms were acquired from BrainWeb (http://www.bic.mni.mcgill.ca/brainweb/) and used to extract cortical surfaces and to generate echo planar imaging (EPI) data. Simulated BOLD signals as the gold standard of activation in our experiment were applied to the surfaces and projected to the volume space with random noise. Three-dimensional isotropic Gaussian kernel smoothing and two-dimensional heat kernel smoothing were applied to the volume- and surface-based methods. Sensitivity and 1-specificity, which are truly and falsely detected activations, and similarity measures, which are spatially and statistically similar for the gold standard and detected activations, were calculated. In the results, the surface-based method showed the sensitivity and similarity scores of about 12% higher than the volume-based method. In conclusion, the surface-based method guarantees better spatial accuracy for the localization of BOLD signal sources within the cortex than the volume-based method.

Keywords: Functional magnetic resonance imaging; Activation detection; Surface-based method; Phantom study; MRI methods/techniques

Introduction

In functional magnetic resonance imaging (fMRI) studies, volume-based methods have been widely used to detect human brain activation (Friston et al., 1995; Acton and Friston, 1998). The topology of the human cortex, having a highly convoluted 2-dimensional structure (Zilles et al., 1988), is seldom considered in the detection procedures of volume-based methods (Kiebel et al., 2000). In most previous studies, however, cortical surface mapping has been used only for visualization purposes and further analyses (e.g., multiple comparisons, group statistics) using the efficiency of surface registration (Van Essen and Drury, 1997; Fischl et al., 1999; Tzourio-Mazoyer et al., 2002; Desai et al., 2005; Argall et al., 2006). To overcome the topological insufficiency of volume-based methods, surface-based methods were proposed using accurate cortical topology in the entire analysis procedure (Andrade et al., 2001).

The volume- and surface-based methods use different procedures to one another (see Fig. 1). Andrade et al. (2001) showed that the two procedures used to analyze fMRI data do not make any statistical difference (namely voxel-versus nodewise statistics) and that the spatial smoothing procedure is the main cause of any topological differences in activation detected by the two methods. Spatial smoothing is one of the most important procedures used to increase the signal-to-noise ratio (SNR) of blood oxygenation level-dependent BOLD signals and to provide statistical smoothness to image data (Worsley, 2005). For volume-based methods, smoothing is usually implemented as an isotropic convolution with a 3-dimensional Gaussian kernel and applied uniformly over the entire intracranial area of volumic data (Cox, 1996). As this leads to averaging of signals emanating from both the cortex and the white matter (WM) or cerebrospinal fluid (CSF), the resultant signal can be influenced not only by sources that are close to each other in a Euclidean sense, but also those that are geodetically distant (e.g., two points, each on opposite sides of a sulcus) (Andrade et al., 2001). To reduce this problem, some improved smoothing techniques have been proposed to account for the

* Corresponding author. Fax: +82 2 2296 5943.
E-mail address: ljm@hanyang.ac.kr (J.-M. Lee).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2006.09.047

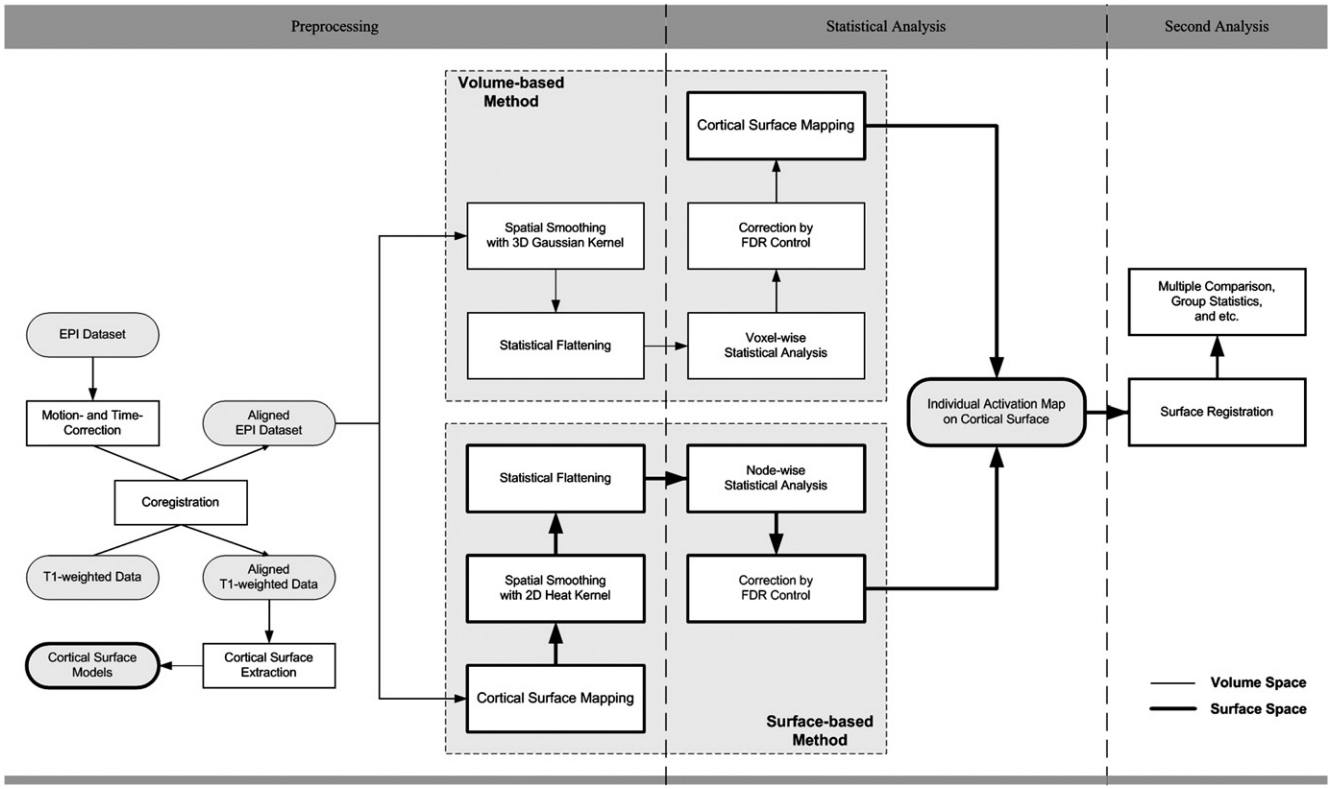


Fig. 1. Procedures for volume- and surface-based fMRI analyses using cortical surface mapping.

anatomical configuration of the brain. For instance, Kiebel et al. (2000) proposed the anatomically informed basis functions that apply 3-dimensional Gaussian kernels with different kernel sizes to the different loci in gray matter (GM). However, for the surface-based method, Chung et al. (2005) proposed more stringent 2-dimensional formed smoothing techniques, such as diffusion smoothing and heat kernel smoothing, that aim to obtain a spatial correlation structure that depends on geodesic rather than Euclidean distances. In this way, the introduction of an artificial correlation between areas that are close together in Euclidean terms but geodetically far apart can be avoided (Chung et al., 2005).

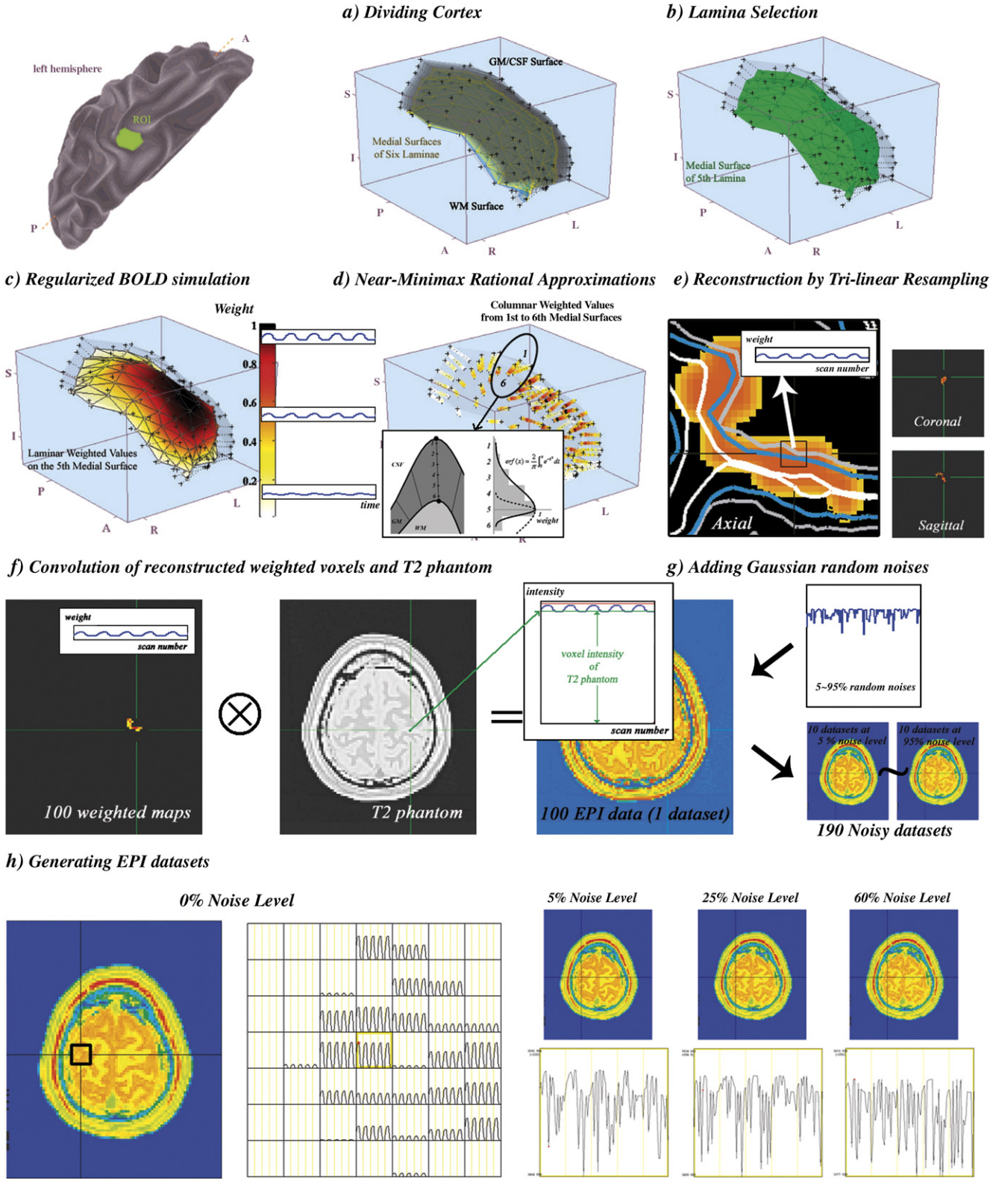
For the precise localization of activation sources using fMRI data, the accurate cortical topology is important prior knowledge and a surface-based method might also be able to provide the topological information through its entire procedure. Currently, the surface-based method for practical applications needs more evaluation, and the accuracy of detecting activation might be fundamentally evaluated in the specialized purpose of fMRI analyses, which is the localization of activation sources. As far as we know, however, there is only one quantitative evaluation that compares the activation detected by each method (Andrade et al., 2001). Andrade et al. (2001) validated volume- and surface-based methods using random noise scans for observing type I errors and real MRI data for demonstrating the local sensitivity of the two methods. Because the main purpose of their evaluation was the validation of the statistical equivalency between the volume- and surface-based methods, the results were insufficient to explain the performance of the two methods in detecting activation in terms of their localization accuracy. Furthermore, it was not clear which of the two results was more similar to the actual loci and extent of activation in the existing evaluation because (i) it was difficult to discover the true neural activation that occurred through their experimental stimuli using MR imaging alone, and (ii) physiological, physical, and procedural confounds may have played a role in real MRI experiments. The limitations of real MRI data have been partially solved by the use of MRI phantoms and simulated BOLD signals in existing studies that focus on validation of several methods for neuroimage analyses and hemodynamic modeling (Collins et al., 1998; Purdon et al., 2001). For instance, Purdon et al. (2001) generated a locally regularized spatiotemporal (LRST) model using a simulated BOLD formulation and compared the fitting performance of statistical models in fMRI analysis. Gautama and Van Hulle (2004) evaluated the optimal degree of spatial smoothing using synthetic fMRI data. Holden et al. (2002) quantified small cerebral ventricular volume changes in growth hormone-treated patients using nonrigid registration with BrainWeb (http://www.bic.mni.mcgill.ca/brainweb/) phantoms. Many more simulation and phantom studies have been conducted, each appropriate for its own goal, and there is a general agreement that simulated BOLD signals and phantoms need to be used for quantitative evaluation to overcome the limitations of real MRI data.

The purpose of this study was to perform a direct quantitative evaluation of the spatial accuracies of fMRI activation detected by the volume- and surface-based methods using MRI phantoms and simulated BOLD signals. In particular, we focus on the effect of the smoothing techniques that were considered as the main difference between the two methods.

Materials and methods

Phantoms and the gold standard of activation

We used BrainWeb to acquire a T1-weighted phantom as the source of cortical surfaces and a T2-weighted phantom as the source of simulated EPI images (Collins et al., 1998; Aubert-Broche et al., 2006). The T1-weighted phantom was acquired from the simulated MRI volumes for the normal brain (1 mm × 1 mm planar pixel size, 1 mm slice thickness, 0% nonuniformity and noise). As it was difficult to directly represent BOLD contrast EPI sequences using BrainWeb, the T2-weighted



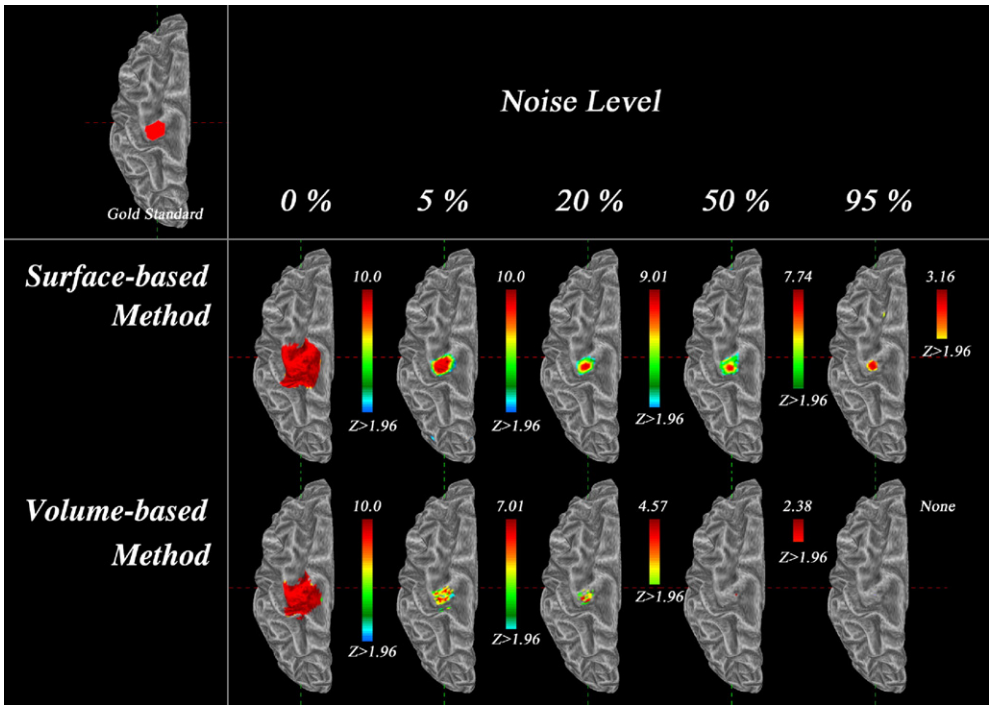


Fig. 3. Activation and its probability distribution detected by volume- and surface-based methods ($P<0.05$, corrected) are visualized on the inflated WM surface.

phantom was acquired from custom MRI simulation (1 mm × 1 mm planar pixel size, 2 mm slice thickness, 0% nonuniformity and noise) with a T2-weighted spin-echo sequence (TR = 2000 ms, TE = 40 ms). As the difference between EPI data acquired from T2-weighted spin-echo, gradient-echo, and T2*-weighted EPI sequences influences only their SNR and contrast-to-noise ratio (Giovagnoni et al., 1995), the difference between the simulated EPI sequences without any artifact seldom affects the topology of imaged brain tissues. The T2-weighted phantoms acquired from the BrainWeb were linearly resampled in a 2 mm isocubic lattice considering the resolution of commonly used real fMRI data (about 1–3 mm).

Two cortical surface models, the WM surface and GM/CSF surface, were extracted from the T1-weighted phantom using the constrained Laplacian-based automated segmentation with proximities algorithm (MacDonald et al., 2000; Kim et al., 2005a). We divided the cortex into six laminae with equidistance between linked nodes on the WM and GM/CSF surfaces and generated six surfaces at the centers of each lamina in order to present parameterized columnar activation in the following procedures (see Fig. 2a).

We manually defined a 2-dimensional region of interest (ROI) in the primary motor area on the medial surface of the fifth lamina from the GM/CSF surface using SUMA software (http://afni.nimh.nih.gov/afni/suma/). We located the BOLD signals on the ROI in the following reconstruction procedure. This ROI was used as the gold standard for true activation in this study (see Fig. 2b).

BOLD simulation and EPI data reconstruction

To simulate a BOLD signal, a simple experimental design with iterative on-and-off stimuli was applied to the BOLD formulation of the LRST model without drift and noise terms (Purdon et al., 2001) (see Fig. 2c). Purdon et al. assumed that the changes in the T2*-based fMRI signal are primarily due to the net average

Fig. 2. Generating EPI data sets. (a) Division of the cortex into six laminae with equidistance between linked nodes on WM and GM/CSF surfaces. (b) Selection of the depth of the activation source (the center of the fifth lamina) and defining a 2-dimensional ROI to be used as the weighted values of the simulated BOLD signal in the primary motor cortex. (c) Regularization of the weighted values of laminar activation. (d) Presentation of the weighted values of columnar activation. The two thick lines indicate the GM/CSF and WM surfaces. The two black dots are the correlated nodes on each surface, and numbers (1–6) are the centers of the laminae divided in panel a. Two thick curves indicate the error functions of the Gaussian distribution, and we used half of each (solid lines). For example, the integral weights of the loci 1–5 are estimated by setting the minimum value to zero and the maximum value as the weighted value of the fifth lamina toward the top of the graph. (e) Reconstruction of the six 2-dimensional weighted values to the 3-dimensional weighted voxels. The colored voxels are the trilinearly resampled weighted values, and the thick lines are the cortical surfaces scanned to this axial intersection. (f) EPI data generated from convolution of panel e and the T2-weighted phantom. (g) Generation of noisy EPI data sets. (h) The axial intersections of the EPI data sets (in the upper row) and the intensity changes in the same location and scan number at each noise level. All data sets are made by convolution of the base image (the T2-weighted phantom of MNI BrainWeb), the reconstructed spatiotemporal weighted voxels, and zero-mean Gaussian random noises. Red dots of lower graphs mean scan numbers. The 'noise level' denotes the maximum magnitude ratio of zero-mean Gaussian random noises for the maximum intensity change of voxels (see Eq. (1)). For instance, the magnitude range of noise is about 0–1 at the 5% noise level and about 0–5 at the 25% noise level where the maximum intensity change is 20. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

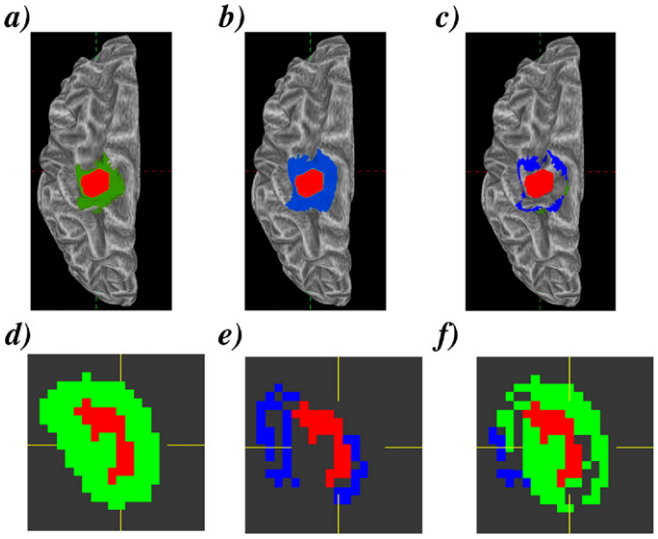


Fig. 4. The spatial influence of smoothing kernels. The gold standard provides 87 nodes (151 voxels), the area blurred by the heat kernel smoothing has 184 nodes (770 voxels), and the area blurred by the 3-dimensional Gaussian kernel smoothing has 124 nodes (2427 voxels). The 2-dimensional heat kernel blurred more nodes on the surface space, and the 3-dimensional Gaussian kernel blurred more voxels in the volume space. Therefore, the equivalent 3-dimensional Gaussian kernel and 2-dimensional heat kernel have different topologies to each other in each space. Panel a shows the nodes blurred by the 3-dimensional Gaussian kernel and mapped from the volume space. Panel b shows the nodes mapped from the volume space and blurred by the 2-dimensional heat kernel. The green and blue areas are blurred nodes from the gold standard (the red area and voxels) by each smoothing technique. The blue and green areas of panel c denote the extra area of the intersection of panels a and b (c = a − b). Furthermore, panel d shows the voxels blurred by the 3-dimensional Gaussian kernel, and panel e shows that the voxels in the gold standard are the sum of all voxels blurred and projected from the six medial surfaces in Fig. 2. Panel f denotes the extra voxels of the intersection of panels d and e. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

deoxyhemoglobin in the voxel, which is the product of the blood volume and the blood oxygenation. The authors also assumed that local blood flow changes and local tissue oxygen consumption changes are tightly coupled in time, though not necessarily proportionally coupled. The deoxygenation term was dictated in terms of oxygen consumption and blood flow by Fick's principle (conservation of oxygen). The BOLD model form was also guided by the empirical observation that the hemodynamic component of the fMRI response for block stimuli has four characteristics: delay, peak response, droop, and poststimulus undershoot. By combining the basic properties of the BOLD signal change with these empirical observations, the authors could create a model of fMRI signal change in response to a stimulus input. In a given pixel, they made the assumption that the blood hemoglobin changes are driven by the stimulus through a linear coupling (2001).

We used the amplitudes of the activation (f_a=1, f_b=−0.5, and f_c=0.2 for flow, volume, and interaction term) and time constants (d_a=1.5 s and d_b=12 s for hemodynamic impulse response and blood volume response). This simulated BOLD signal was spatially regularized by applying a finite impulse-response Epanechnikov weighting kernel to the gold standard to present parameterized laminar activation (Purdon et al., 2001). As all nodes on the WM surface have corresponding nodes on the GM/CSF surface (Kabani et al., 2001), we could present columnar activations distributed from the fifth lamina to the first and sixth laminae using the near-minimax rational approximations for the error function, which were twice the integral of the zero-mean Gaussian distribution (Cody, 1969). All values mapped to the nodes of the laminar surfaces were the weighted values given by the simulated BOLD signal (see Fig. 2d).

We projected the weighted values in the surface space to the trilinearly resampled weighted voxels in the space of the T2-weighted phantom (see Fig. 2e). Three methods were considerable as the projection method mapping surface data to the voxel space. They were nearest neighborhood (NN), trilinear (TL) and tricubic (TC) interpolations. In our preliminary experiment for this projection procedure, NN, TL, and TC showed the different maximum intensity values (4251, 4357, and 4353 each) for maximum BOLD signal power in our ROI each, but those did not show significant difference. We selected the TL interpolation because it showed larger BOLD signal changes than the others.

Then, we generated EPI data convoluting the T2-weighted phantom and the reconstructed weighted voxels (see Fig. 2f). As the BOLD signal change associated with functional activation is 1–4% for voxel intensity change (Jezzard and Clare, 1999), the maximum intensity change occurring from the weighted voxels was limited to 4% for the intensity of each voxel in the T2-

Table 1
The sensitivity (TPF), one minus specificity (FPF), and probabilistic similarity measures of activations detected by the volume- and surface-based methods at the 0% noise level

Measure	Volume-based method	Surface-based method
TPF	1	1
FPF	0.0104	0.0158
PSI	0.2908	0.2129
POF	1	1
PEF	4.8761	7.3903

weighted phantom. One hundred time-series EPI data were generated and defined as the data set at 0% noise level.

Noise in the fMRI data was assumed to be zero-mean Gaussian random ($\Delta I_{N(0,\sigma_n)}$) and uncorrelated temporally and spatially in this study. We generated noisy EPI data sets by adding noise to the data set at 0% noise level. The noise level is defined as:

$$\eta(\%) = \frac{\max\{\Delta I_{N(0,\sigma_n)}\}}{\max\{\Delta I_{\mathrm{BOLD}}\}} \times 100 \quad (1)$$

where $\Delta I_{N(0,\sigma_n)}$ and ΔI_{BOLD} denote intensity changes by the noise, N, and by the simulated BOLD signal in each EPI data set. The noise was applied from a 5% to 95% level (19 levels, increasing 5% per level). Ten sets of the random noise were added to one data set at each noise level. Ultimately, 180 EPI data sets were generated (see Figs. 2g and h).

Data processing

The entire procedure of the volume- and surface-based methods followed the study of Andrade et al. (2001). However, we skipped alignment procedures between the T1-weighted phantom and EPI data sets for motion and time correction of EPI data because BrainWeb phantoms were already spatially matched. Statistical flattening was also skipped because the second analysis (e.g., multiple comparisons and group statistics) was not included in this study. The orders of spatial smoothing procedures and cortical surface mapping procedures in the volume- and surface-based methods are different because the spaces of each are completely

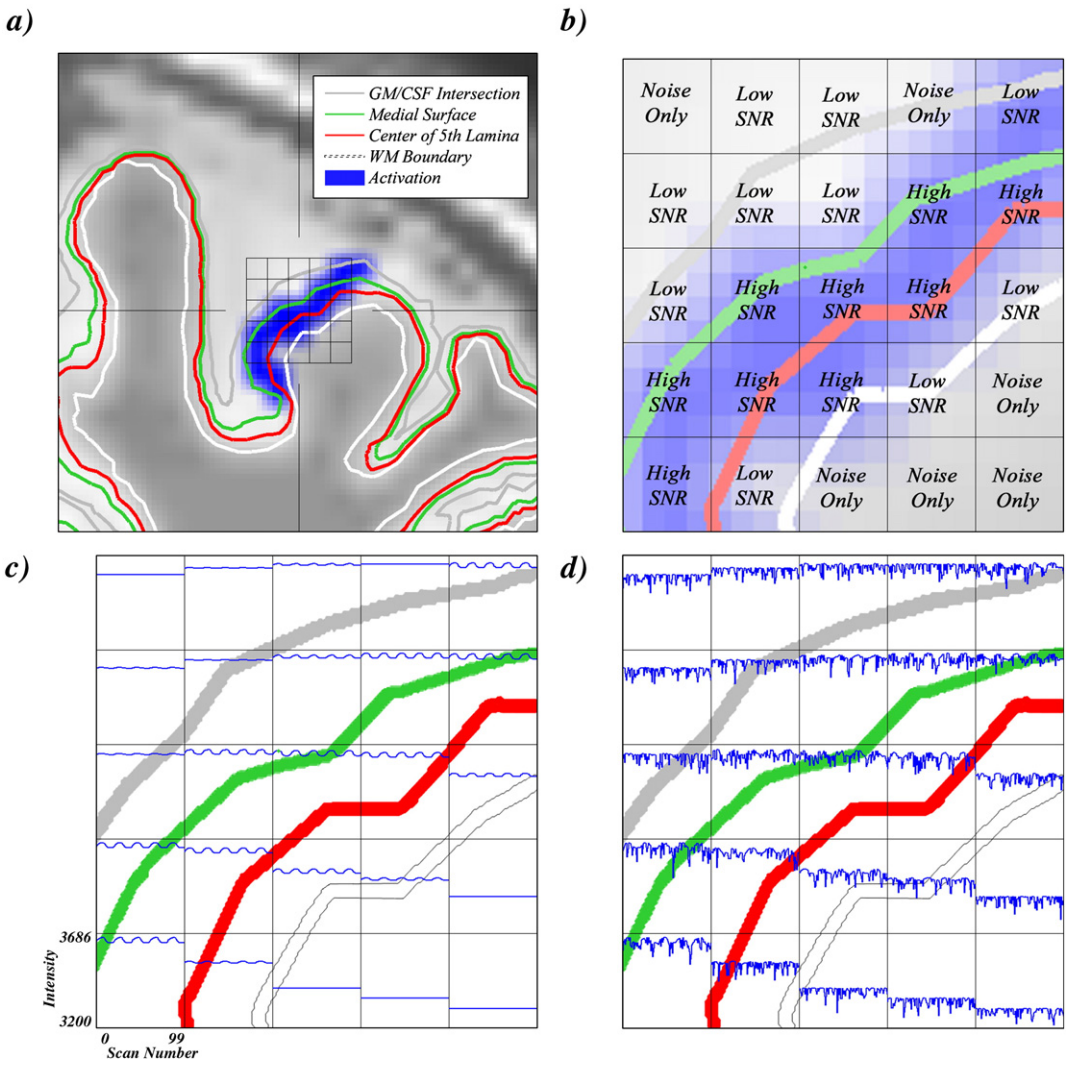


Fig. 5. The true BOLD signals and noise near the gold standard. (a) A coronal intersection of the activated region in EPI data. The blue area was activated by the simulated BOLD signals. The gray, green, red, and white thick lines, scanned to the voxel space, denote the GM/CSF surface, the medial surface between the GM/CSF and WM surfaces, the medial surface of the fifth lamina, and each WM surface. (b) The SNR of the EPI data. (c) The black thin grids in the center of the figure are voxel boundaries. Each grid is located at the same position shown in the center of panel a. The blue lines denote the time course of the true BOLD signals at each voxel, and they indicate intensity changes in the EPI data at the 0% noise level. (d) The blue lines denote the intensity changes of the EPI data at the 5% noise level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

different, and the details are mentioned in the following explanation of two methods.

For the volume-based analysis, all EPI data were blurred using a 3-dimensional isotropic Gaussian kernel with 4 mm full-width at half-maximum (FWHM). Then, the intracranial areas of all blurred EPI data were automatically masked and analyzed using the AFNI package (http://afni.nimh.nih.gov/afni). Activation was detected voxelwise by measuring cross-correlation coefficients (Saad et al., 2001) between each data set and a gamma variate function (Cohen, 1997) as the ideal reference of hemodynamic response function. Detected activation was corrected by a false discovery rate controlling procedure (Genovese et al., 2002), and the same threshold ($P<0.05$) was applied to all activation maps. These voxelwise statistical results were assigned to the cortical surface using a center-of-normal (CN) approach. We applied nearest neighborhood sampling to each mapping approach to remove the effects of assigning kernel size and interpolation type (Desai et al., 2005).

For the surface-based method, it is necessary first to assign EPI data to the nodes of the cortical surface before the smoothing procedure, and we evenly applied the CN approach used in the volume-based method. After the cortical surface mapping procedure, we applied heat kernel smoothing with an equivalent bandwidth to the 3-dimensional Gaussian kernel ($\sigma=1.6986$, 11 iterations) that was derived from the relationship between FWHM and the bandwidth (FWHM $=\sqrt{8\ln 2}\sigma$) (Chung et al., 2005). Detection and correction of activation were made nodewise, implemented by the same procedures used in the volume-based method.

Similarity measures

We compared the activation detected by volume- and surface-based methods with the gold standard for calculating sensitivity and specificity (Anbeek et al., 2005). True positives (TP) and true negatives (TN) are the numbers of correctly detected and undetected nodes, and the false positives (FP) and the false negatives (FN) are the numbers of falsely detected and undetected nodes. The true positive fraction (TPF), that is, the sensitivity, and the false positive fraction (FPF), that is, one minus specificity, were calculated for the different noise levels from 0 to 95%.

$$\mathrm{TPF}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{2}$$

$$\mathrm{FPF}=\frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}} \tag{3}$$

We also calculated the probabilistic similarity index (PSI), probabilistic overlap fraction (POF), and probabilistic extra fraction (PEF) to observe the spatial and statistical influences of the smoothing techniques in both methods (Anbeek et al., 2005):

$$\mathrm{PSI}=\frac{\Sigma P_{x,\mathrm{gs}=1}}{\Sigma 1_{x,\mathrm{gs}=1}+\Sigma P_x} \tag{4}$$

$$\mathrm{POF}=\frac{\Sigma P_{x,\mathrm{gs}=1}}{\Sigma 1_{x,\mathrm{gs}=1}} \tag{5}$$

$$\mathrm{PEF}=\frac{\Sigma P_{x,\mathrm{gs}=0}}{\Sigma 1_{x,\mathrm{gs}=1}} \tag{6}$$

where $\Sigma P_{x,\ \mathrm{gs}=1}$ of Eqs (4) and (5) is sum over the probability scores of all detected activation nodes for the gold standard, $\Sigma P_{x,\ \mathrm{gs}=0}$ of Eq. (6) is sum over the probability scores of all detected activation nodes outside of the gold standard, $\Sigma P_{x,\ \mathrm{gs}=1}$ is the number of all nodes for the gold standard, and ΣP_x of Eq. (4) is sum over the probabilities of all detected activation nodes. The PSI and POF have values from 0 to 1, and higher PSI values indicate that the loci and extent of activation are closer to those of the gold standard with stronger probabilities.

We performed a Wilcoxon test for the PSI scores of the volume-based method with the surface-based method.

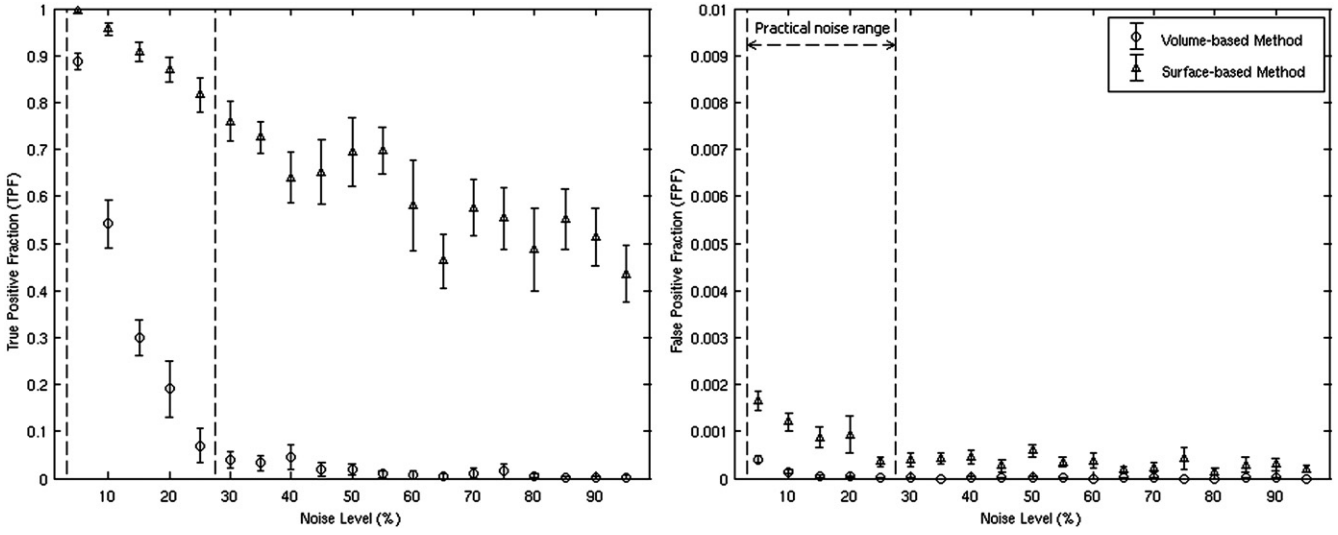


Fig. 6. The means and standard deviations of the TPF scores, which are the sensitivities at overall noise levels (left graph), and the FPF scores, which are the 1-specificities over 5% noise levels. The practical noise range is indicated by the two dotted lines.

Kernel size effect

To compare the effect of smoothing kernel size, we repeated data processing with different smoothing kernel sizes (6 mm, 8 mm, and 10 mm FWHM). We performed Wilcoxon tests for the PSI scores of the volume- with the surface-based method for each smoothing kernel size, and Kruskal-Wallis tests for the PSI scores between the different kernel sizes at each nose level.

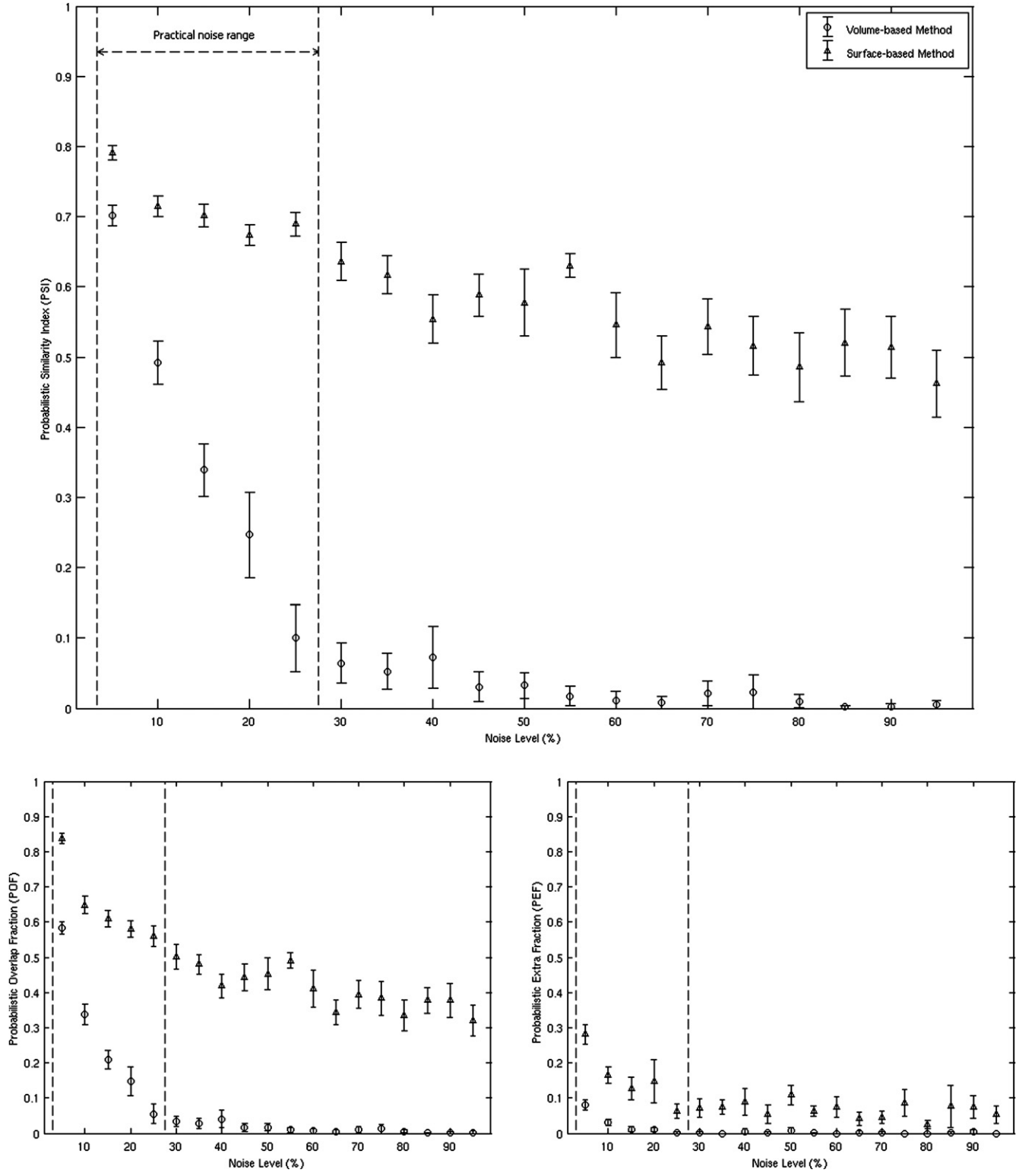


Fig. 7. The means and standard deviations of the probabilistic similarity measures. The large upper graph shows the mean PSI scores and standard deviations. The lower left graph shows the POF scores, and the remainder shows the PEF scores over 5% noise levels. The practical noise range is marked by the two dotted lines.

Results

Activation maps

The probability of activation detected by the surface-based method was higher than that of the volume-based method at every noise level. Furthermore, the peaks of activation detected by the surface-based method were distributed more regularly in the region of the gold standard (see Fig. 3). Little significant activation was detected in the results of the volume-based method at over 50% noise levels ($P<0.05$, corrected).

Effect of the smoothing kernels

The resultant activation maps at the 0% noise level showed the spatial range of the gold standard alone blurred and expanded by each smoothing kernel in the surface space. The 2-dimensional heat kernel and 3-dimensional Gaussian kernel, which had equivalent bandwidths in each native space, did not have equivalent topology as Fig. 4 shows. The 2-dimensional heat kernel blurred more nodes (about 68%) in the surface space. At the 0% noise level, extra nodes blurred by the smoothing kernels were partially synchronized to the BOLD signals for the gold standard, and this effect influenced the FPF and PEF involving the FP error. Table 1 shows the similarity measures at 0% noise levels. The surface-based method covered more extra nodes than the volume-based method (about 50% greater FPF and PEF). Where the TPF and POF of both had the same value (=1), the PSI of the volume-based method seemed to produce better scores than that of the surface-based method because there are relatively fewer FP errors in the volume-based method.

Similarity measures of the noisy data sets

The results at the 0% noise level were a reversal over 5% noise levels because the 3-dimensional kernel included more noisy signals on the outside of the cortex than the 2-dimensional kernel (see Figs. 4 and 5). This effect was presented in the trend of the similarity measures. As stronger noise was added to the EPI data, all measures of the volume-based method tend toward low scores more rapidly than those of the surface-based method. Fig. 6 shows the mean TPF and FPF with their standard deviations at noise levels ranging from 5% to 95%. The surface-based method was more sensitive because the mean TPF of the surface-based method shows higher values of 12.45% more than the volume-based method at all noise levels. Both methods showed similar FPF scores of less than 0.35% (0–142 nodes) for the whole TN area (40,875 nodes) at all noise levels. The specificities of both methods are guaranteed by their much lower FPF scores. Fig. 7 shows the similarity measures for each method over 5% noise levels. The mean PSI scores and the mean POF scores of the surface-based method were 12.74% and 43.37% higher than those of the volume-based method. The mean PEF scores of the surface-based method were higher under 250.01% than those of the volume-based method, but the PEF scores of both methods approached zero. The differences of the PSI scores were statistically significant through the range of noise levels (see Table 2).

Kernel size effect

The probabilities of activations detected by the surface-based method were higher than those of the volume-based method at every smoothing kernel size (see Figs. 8 and 9). Due to the increase of smoothing kernel size, the means and standard deviations of the similarity measures were increased (see Figs. 10 and 11). The differences of the PSI scores were statistically significant through the range of noise levels for all smoothing kernel sizes (see Table 2). The differences of the smoothing kernel sizes of volume-based method were more significant than those of surface-based method through the range of practical noise levels (see Table 3).

Table 2
Mann–Whitney–Wilcoxon U test results (p-value, double sided significance) testing for PSI differences between volume- and surface-based methods

Noise level (%)	Smoothing kernel size (FWHM, mm)			
	4	6	8	10
5	0.0002	0.0002	0.0002	0.0002
10	0.0002	0.0004	0.0004	0.0002
15	0.0002	0.0004	0.0017	0.0006
20	0.0002	0.0002	0.0002	0.0003
25	0.0002	0.0002	0.0002	0.0002
30	0.0002	0.0002	0.0002	0.0003
35	0.0002	0.0002	0.0002	0.0002
40	0.0002	0.0002	0.0003	0.0004
45	0.0001	0.0002	0.0002	0.0017
50	0.0002	0.0002	0.0002	0.0008
55	0.0001	0.0001	0.0002	0.0002
60	0.0001	0.0002	0.0002	0.0004
65	0.0001	0.0001	0.0003	0.0069
70	0.0001	0.0001	0.0002	0.0005
75	0.0001	0.0001	0.0001	0.0001
80	0.0001	0.0001	0.0001	0.0001
85	0.0001	0.0001	0.0001	0.0001
90	0.0001	0.0001	0.0001	0.0002
95	0.0001	0.0001	0.0002	0.0002

Discussion

Topological and statistical influences of the smoothing kernels

The statistical inferences of the surface-based method are conceptually similar to those of the volume-based method (Andrade et al., 2001). Therefore, the difference between the two methods is mainly due to the smoothing techniques resulting from differences in their geometry. As the volume-based method showed lower similarity scores under noisy conditions, we could therefore conclude that the 3-dimensional Gaussian kernel more profoundly decreased the magnitudes of true BOLD signals by including noise outside of the GM bands than the 2-dimensional heat kernel. This conclusion partially corresponds with existing studies that attempt to apply cortical topology to fMRI analyses (Kiebel et al., 2000; Andrade et al., 2001). Combined with this effect of smoothing kernels, this difference would influence the probability of detected activation as a statistical result of fMRI analysis. As PSI is a measure that shows the probabilistic trend of detected activation, the surface-based method assigned higher probability to the same activation than the volume-based method. As Fig. 3 shows, more activated nodes with higher probabilities were detected in the surface-based method than in the volume-based method at the same threshold ($P<0.05$). Therefore, we could conclude that the surface-based method is able to detect a probability closer to the gold

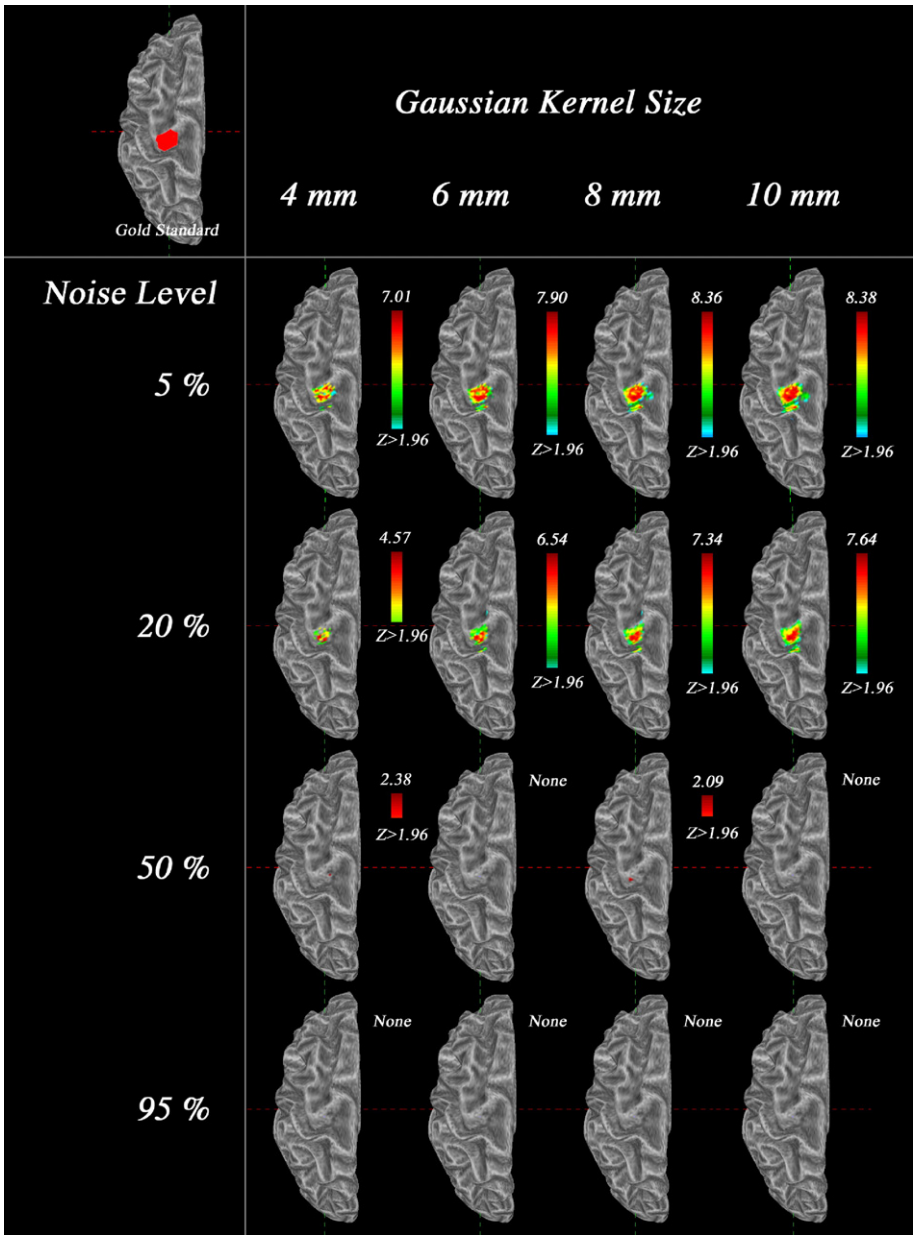


Fig. 8. Activations and their probability distributions detected by volume-based methods with 4, 6, 8, and 10 mm FWHM smoothing kernels ($P<0.05$, corrected) are visualized on the inflated WM surface.

standard (=1) and avoid the effect of strong noise better than the volume-based method.

The effect of different smoothing kernel sizes

An increase in the FWHM of smoothing kernel increased the means and standard deviations of the similarity measures in both volume- and surface-based methods (see Figs. 10 and 11). However, the reproducibility of the surface-based method might be better than the volume-based method since the surface-based method showed less increment in the standard deviations of the measures due to the increase of smoothing kernel size than the volume-based method. As Table 3 shows, smoothing kernel sizes more significantly influenced the statistical results in the volume-based method than the surface-based method through the practical noise levels. Therefore, we could conclude that the activation detected by the surface-based method is less influenced by the FWHM of smoothing kernel than those of the volume-based method.

Several practical issues

According to the existing studies that focus on the SNR difference between practical EPI sequences, the intensities of EPI data mainly affected by the nongeometric artifacts (e.g., local field homogeneity and noise) show large variations in the range of 6% to 23% (Hillenbrand et al., 1999). Although the PSI differences over 30% noise levels are considered to be practically extreme cases, in our study, the surface-based method showed even higher PSI scores under

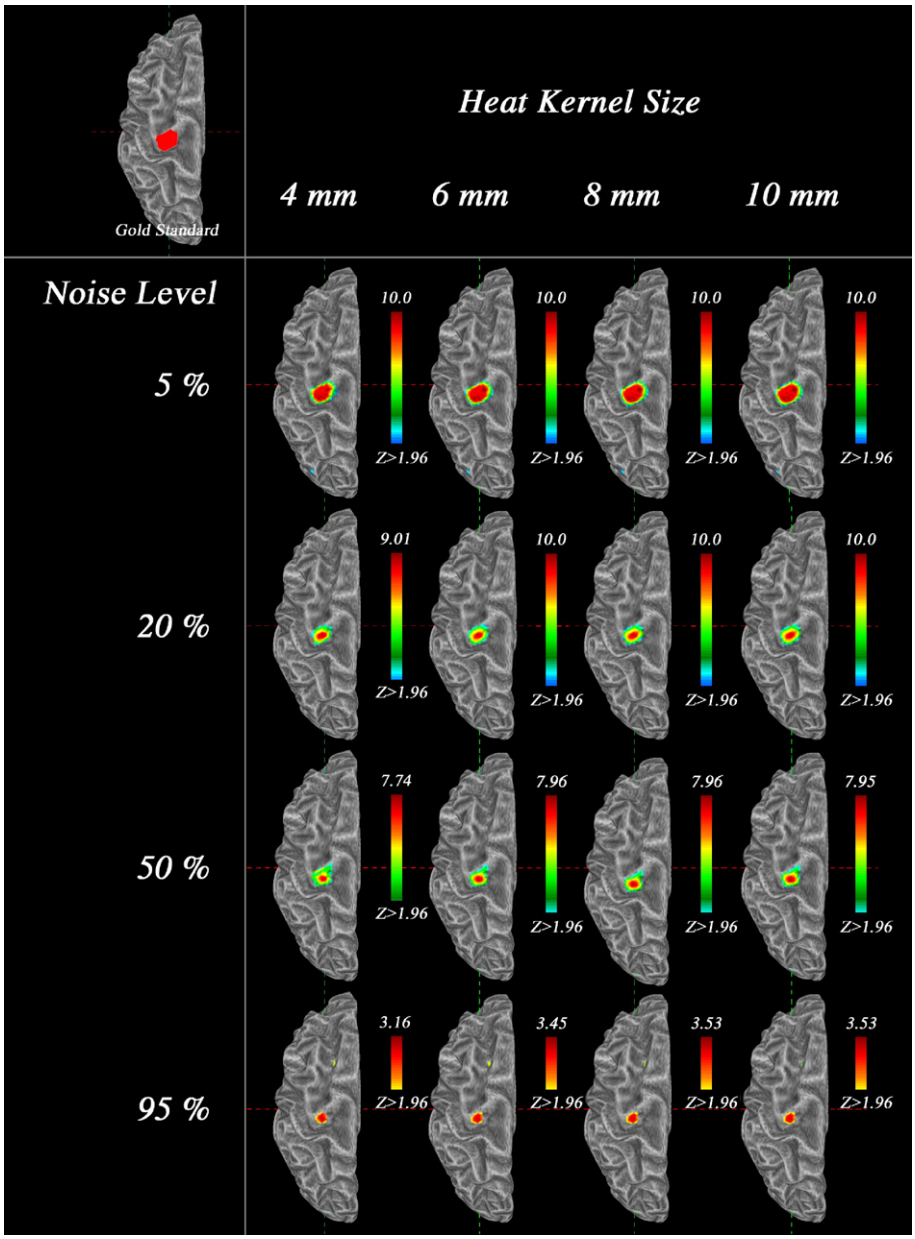


Fig. 9. Activations and their probability distributions detected by surface-based methods with 4, 6, 8, and 10 mm FWHM smoothing kernels ($P<0.05$, corrected) are visualized on the inflated WM surface.

25% noise levels than did the volume-based method. The differences in brain tissue intensity (GM versus CSF or WM) could also decrease the magnitude of the true BOLD signal in 3-dimensional Gaussian kernel smoothing. This partially corresponds to the existing report that volume-based smoothing increases the partial volume effect near a narrow GM band, and it consequently decreases the probability of detected activation in PET analysis (Frouin et al., 2002). This may be evidence that the difference between imaged tissue contrasts led by different EPI sequences is one of the main factors influencing the degree of BOLD signal reduction in volume-based smoothing procedures. In the case that cortical surfaces and EPI data are spatially matched, surface-based smoothing, in contrast, could not increase the partial volume effect because it is implemented in GM bands. Therefore, the surface-based method could be seldom influenced by the variation of imaged tissue contrasts. The PSI difference between the volume- and surface-based methods might consequently exist in almost all cases although the EPI sequence is changed.

Another substantial issue leading to the different results from the two methods is the cortical surface mapping used to convert data from the volume space to the surface space. There are two frequently used cortical surface mapping methods. The first is the CN method used in FreeSurfer (http://surfer.nmr.mgh.harvard.edu/). The medial surface is produced from the GM/CSF and WM surfaces, and the volumetric data crossing to this medial surface are assigned. The second is the two-surface method used in SUMA (http://afni.nimh.nih.gov/afni/suma). It uses the GM/CSF and WM surfaces together and assigns values as the greater absolute value between the values assigned to the two surfaces. The common issue of these methods is that the original signals can potentially be lost or smoothed in a variety of ways during the transfer, directly affecting

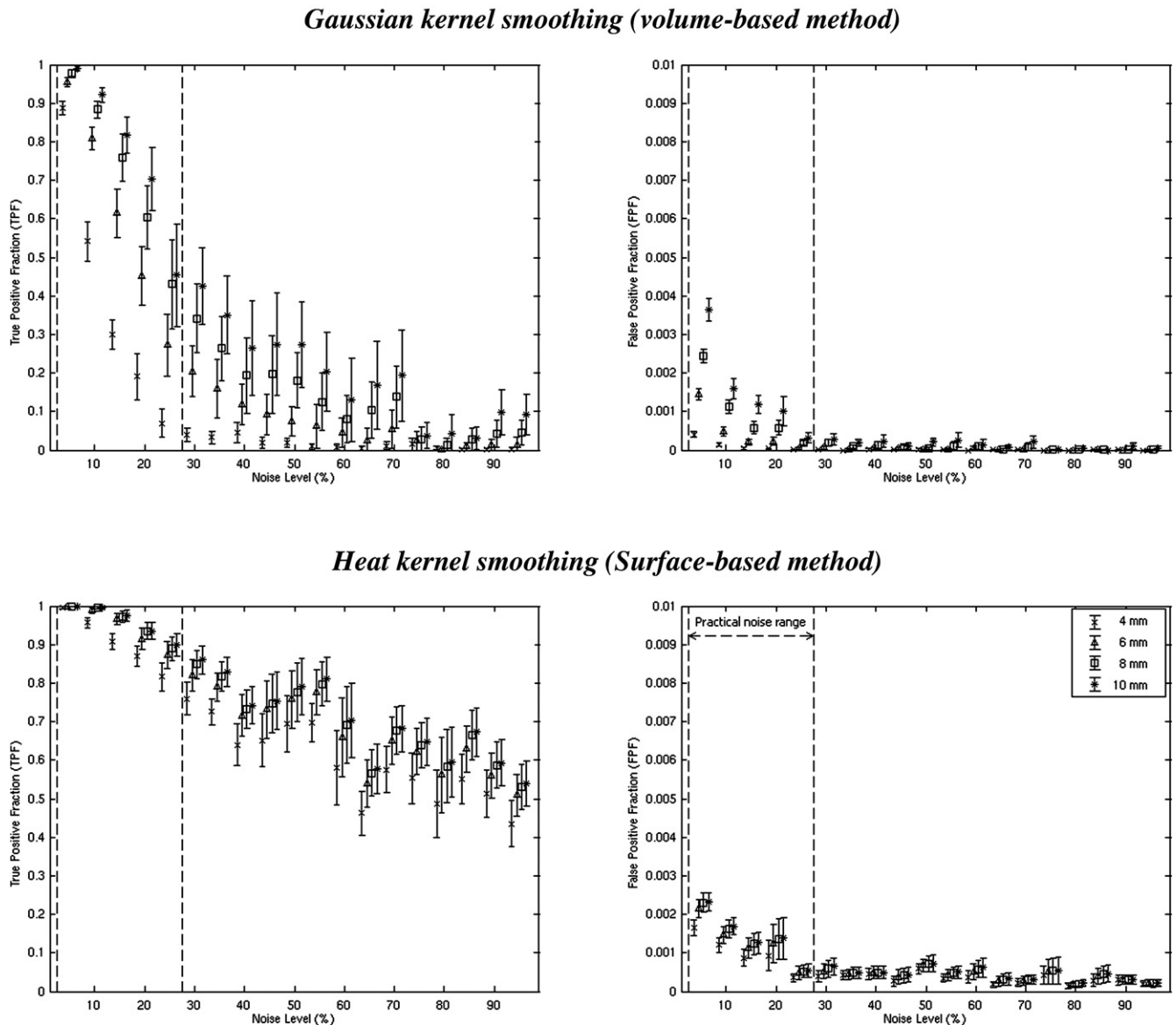


Fig. 10. The means and standard deviations of the TPF and FPF scores of volume- and surface-based methods with 4, 6, 8, and 10 mm FWHM smoothing kernels.

the accuracy of the results as functional activation along the entire thickness of GM is mapped to a single node on the surface (Desai et al., 2005). However, Saad et al. (2005) mentioned that this might not influence the resultant activation map because fMRI data are currently acquired with low resolution by commonly used 1.5 or 3 T MRI devices.

Existing limitations

Currently, the advantages of cortical surface models and 2-dimensional smoothing techniques seldom influence the accuracy of detected activation in practical applications because the voxel intensity of EPI data is an averaged value of signals generated in an arbitrary cubical area, and the clinically useful EPI data for whole brain imaging have low spatial resolution (Liang et al., 2000). With the improvement of high-field MRI devices, however, these problems could be solved by ultrahigh resolution MRI data (Harel et al., 2006; Zhao et al., 2006). Furthermore, both the volume- and the surface-based methods are not applicable for detecting the activation of cerebellar and subcortical structures at this time. This issue might be addressed by modeling techniques for small and complex brain structures (Kim et al., 2005b).

Because we used a simplified model focusing on only the spatial extent of the activation influenced by the topological effects of the smoothing techniques, there are some limitations of this study in explaining other effects of the smoothing. For instance, we neglected the temporal terms (e.g., hemodynamic delays and temporal autocorrelations) and generated the simple columnar activations without considering the physiological properties of different neuron types at every cortical depth. However, we expect that more realistic evaluation of the volume- and the surface-based methods could be performed by modeling the previous findings via cytoarchitectural studies showing relatively accurate properties corresponding to the nature of columnar and laminar activation

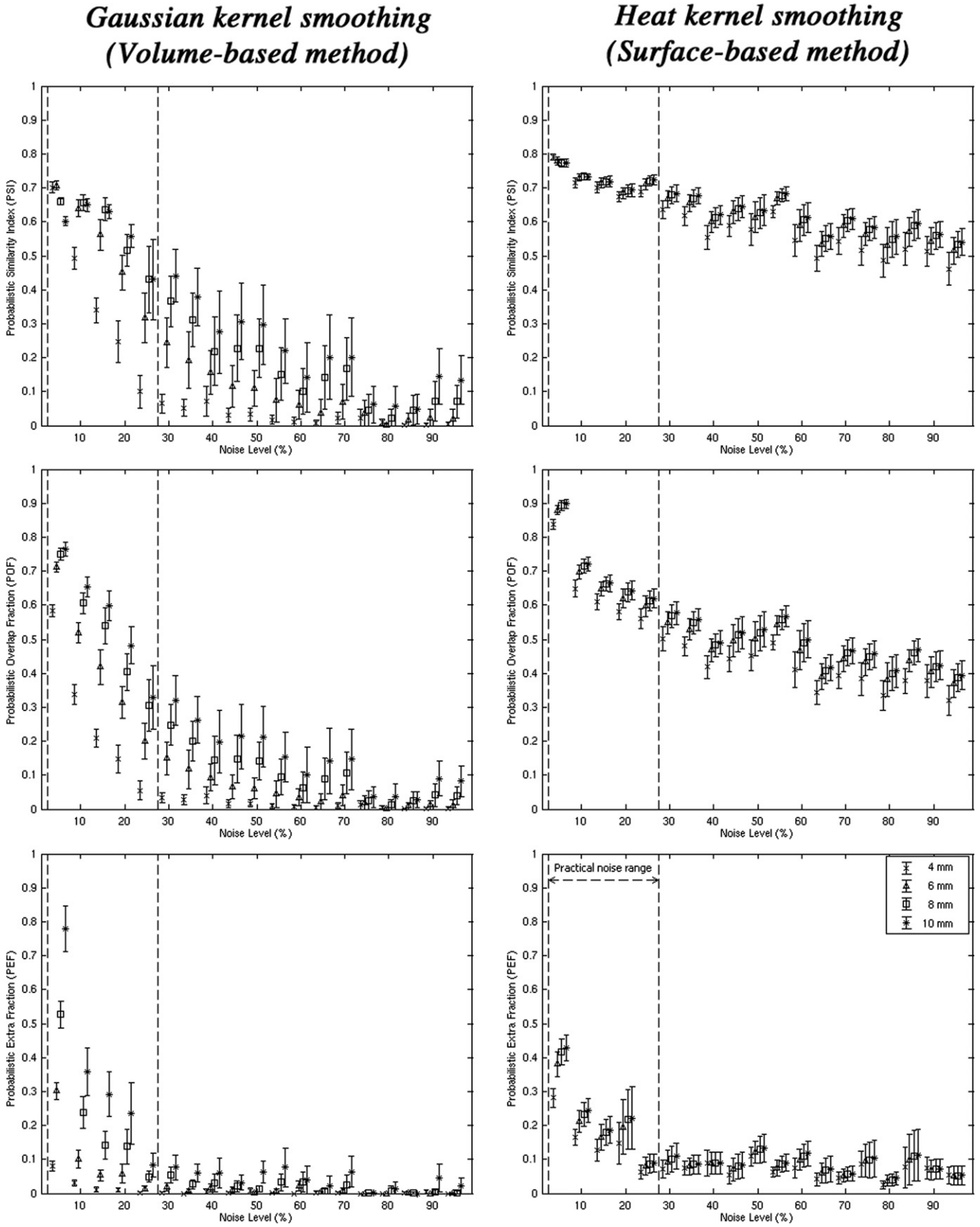


Fig. 11. The means and standard deviations of the probabilistic similarity measures of volume- and surface-based methods with 4, 6, 8, and 10 mm FWHM smoothing kernels.

Table 3
Kruskal–Wallis test results (p-value) testing for PSI differences between smoothing kernel sizes

Noise level (%)	Volume-based method	Surface-based method
5	0.0001	0.2339
10	0.0001	0.1626
15	0.0001	0.4442
20	0.0001	0.4505
25	0.0022	0.0820
30	0.0001	0.2092
35	0.0007	0.0457
40	0.2130	0.1092
45	0.0038	0.1726
50	0.0312	0.5309
55	0.0443	0.0070
60	0.2454	0.2695
65	0.1551	0.1783
70	0.4026	0.1529
75	0.8991	0.0534
80	0.7589	0.2621
85	0.5504	0.1817
90	0.1070	0.5223
95	0.0309	0.1288

(Koenderink and Uylings, 1996; Lubke et al., 2003; Chance et al., 2004).

Conclusion

In this work, we have used spatially matched phantoms, simulated BOLD signals, and controlled random noise to quantitatively evaluate the performance of volume- and surface-based methods for fMRI analysis. Measuring the probabilistic similarity indices, we have shown that the topological difference between the 3- and 2-dimensional smoothing kernels of each method leads to a statistical difference in resultant activation. More generally, the surface-based method for fMRI analysis guarantees better sensitivity and spatial accuracy for the localization of BOLD signal sources within the cortex than the volume-based method.

Acknowledgments

This work was supported by grant No. M10530010002-06N3001-00210 from MOCIE (Korean Ministry of Commerce, Industry and Energy).

References

Acton, P.D., Friston, K.J., 1998. Statistical parametric mapping in functional neuroimaging: beyond PET and fMRI activation studies. Eur. J. Nucl. Med. 25 (7), 663–667.

Anbeek, P., Vincken, K.L., van Bochove, G.S., van Osch, M.J., van der Grond, J., 2005. Probabilistic segmentation of brain tissue in MR imaging. Neuroimage 27 (4), 795–804.

Andrade, A., Kherif, F., Mangin, J.F., Worsley, K.J., Paradis, A.L., Simon, O., Dehaene, S., Le Bihan, D., Poline, J.B., 2001. Detection of fMRI activation using cortical surface mapping. Hum. Brain Mapp. 12 (2), 79–93.

Argall, B.D., Saad, Z.S., Beauchamp, M.S., 2006. Simplified intersubject averaging on the cortical surface using SUMA. Hum. Brain Mapp. 27 (1), 14–27.

Aubert-Broche, B., Evans, A.C., Collins, L., 2006. A new improved version of the realistic digital brain phantom. Neuroimage 32 (1), 138–145.

Chance, S.A., Tzotzoli, P.M., Vitelli, A., Esiri, M.M., Crow, T.J., 2004. The cytoarchitecture of sulcal folding in Heschl's sulcus and the temporal cortex in the normal brain and schizophrenia: lamina thickness and cell density. Neurosci. Lett. 367 (3), 384–388.

Chung, M.K., Robbins, S.M., Dalton, K.M., Davidson, R.J., Alexander, A.L., Evans, A.C., 2005. Cortical thickness analysis in autism with heat kernel smoothing. Neuroimage 25 (4), 1256–1265.

Cody, W.J., 1969. Rational chebyshev approximations for the error function. Math. Comp. 22, 631–638.

Cohen, M.S., 1997. Parametric analysis of fMRI data using linear systems methods. Neuroimage 6 (2), 93–103.

Collins, D.L., Zijdenbos, A.P., Kollokian, V., Sled, J.G., Kabani, N.J., Holmes, C.J., Evans, A.C., 1998. Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imaging 17 (3), 463–468.

Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical Research 29 (3), 162–173.

Desai, R., Liebenthal, E., Possing, E.T., Waldron, E., Binder, J.R., 2005. Volumetric vs. surface-based alignment for localization of auditory cortex activation. Neuroimage 26 (4), 1019–1029.

Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage 9 (2), 195–207.

Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C., Frackowiak, R.S., Turner, R., 1995. Analysis of fMRI time-series revisited. Neuroimage 2 (1), 45–53.

Frouin, V., Comtat, C., Reilhac, A., Gregoire, M.C., 2002. Correction of partial-volume effect for PET striatal imaging: fast implementation and study of robustness. J. Nucl. Med. 43 (12), 1715–1726.

Gautama, T., Van Hulle, M.M., 2004. Optimal spatial regularisation of autocorrelation estimates in fMRI analysis. Neuroimage 23 (3), 1203–1216.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15 (4), 870–878.

Giovagnoni, A., Paci, E., Ercolani, P., Kiefer, B., Santino, P., Piga, A., 1995. MR imaging of focal liver lesions: comparison between turbo spin-echo and conventional spin-echo pulse sequences. Magma 3 (3–4), 143–150.

Harel, N., Lin, J., Moeller, S., Ugurbil, K., Yacoub, E., 2006. Combined imaging–histological study of cortical laminar specificity of fMRI signals. Neuroimage 29 (3), 879–887.

Hillenbrand, C., Deichmann, R., Hahn, D., Haase, A., 1999. Signal intensities in FLASH-EPI-hybrid sequences. J. Magn. Reson. 139 (1), 74–80.

Holden, M., Schnabel, J.A., Hill, D.L., 2002. Quantification of small cerebral ventricular volume changes in treated growth hormone patients using nonrigid registration. IEEE Trans. Med. Imaging 21 (10), 1292–1301.

Jezzard, P., Clare, S., 1999. Sources of distortion in functional MRI data. Hum. Brain Mapp. 8 (2–3), 80–85.

Kabani, N., Le Goualher, G., MacDonald, D., Evans, A.C., 2001. Measurement of cortical thickness using an automated 3-D algorithm: a validation study. Neuroimage 13 (2), 375–380.

Kiebel, S.J., Goebel, R., Friston, K.J., 2000. Anatomically informed basis functions. Neuroimage 11 (6 Pt 1), 656–667.

Kim, J.S., Singh, V., Lee, J.K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J.M., Kim, S.I., Evans, A.C., 2005a. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. Neuroimage 27 (1), 210–221.

Kim, S.H., Lee, J.M., Kim, H.P., Jang, D.P., Shin, Y.W., Ha, T.H., Kim, J.J., Kim, I.Y., Kwon, J.S., Kim, S.I., 2005b. Asymmetry analysis of deformable hippocampal model using the principal component in schizophrenia. Hum. Brain Mapp. 25 (4), 361–369.

Koenderink, M.J., Uylings, H.B., 1996. Morphometric dendritic field analysis of pyramidal neurons in the human prefrontal cortex: relation to section thickness. J. Neurosci. Methods 64 (1), 115–122.

Liang, Z.-P., Lauterbur, P.C., Society, I.E.i.M.a.B., 2000. Principles of Magnetic Resonance Imaging: A Signal Processing Perspective. SPIE Optical Engineering Press, New York, pp. 187–212.

Lubke, J., Roth, A., Feldmeyer, D., Sakmann, B., 2003. Morphometric analysis of the columnar innervation domain of neurons connecting layer 4 and layer 2/3 of juvenile rat barrel cortex. Cereb. Cortex 13 (10), 1051–1063.

MacDonald, D., Kabani, N., Avis, D., Evans, A.C., 2000. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. Neuroimage 12 (3), 340–356.

Purdon, P.L., Solo, V., Weisskoff, R.M., Brown, E.N., 2001. Locally regularized spatiotemporal modeling and model comparison for functional MRI. Neuroimage 14 (4), 912–923.

Saad, Z.S., Ropella, K.M., Cox, R.W., DeYoe, E.A., 2001. Analysis and use of FMRI response delays. Hum. Brain Mapp. 13 (2), 74–93.

Saad, Z.S., Otwinowski, J., Cox, R.W., 2005. Sampling requirements of highly convoluted brain surfaces with fMRI. Proceedings of the 11th Annual Meeting of the Organization For Human Brain Mapping, Toronto, Canada.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15 (1), 273–289.

Van Essen, D.C., Drury, H.A., 1997. Structural and functional analyses of human cerebral cortex using a surface-based atlas. J. Neurosci. 17 (18), 7079–7102.

Worsley, K.J., 2005. Spatial smoothing of autocorrelations to control the degrees of freedom in fMRI analysis. Neuroimage 26 (2), 635–641.

Zhao, F., Wang, P., Hendrich, K., Ugurbil, K., Kim, S.G., 2006. Cortical layer-dependent BOLD and CBV responses measured by spin-echo and gradient-echo fMRI: insights into hemodynamic regulation. Neuroimage 30 (4), 1149–1160.

Zilles, K., Armstrong, E., Schleicher, A., Kretschmann, H.J., 1988. The human pattern of gyrification in the cerebral cortex. Anat. Embryol. 179 (2), 173–179.