Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computer Vision and Image Understanding 99 (2005) 435–452
Computer Vision and Image Understanding
www.elsevier.com/locate/cviu

Robust anisotropic diffusion to produce enhanced statistical parametric map from noisy fMRI

Hae Yong Kim [a,*], Javier Giacomantone [a], Zang Hee Cho [b]

[a] *Escola Politécnica, Universidade de São Paulo, Av. Prof. Luciano Gualberto, trav. 3, 158, CEP 05508-900, São Paulo, SP, Brazil*
[b] *Department of Radiological Sciences, University of California at Irvine, Medical Sciences I, Room B140, 92697-5000, Irvine, CA, USA*

Received 22 April 2003; accepted 12 April 2005
Available online 9 June 2005

Abstract

This paper presents a new, simple, and elegant technique to obtain enhanced statistical parametric maps (SPMs) from noisy functional magnetic resonance imaging (fMRI) data. This technique is based on the robust anisotropic diffusion (RAD), a technique normally used as an edge-preserving filter. A direct application of the RAD to the fMRI data does not work, because in this case RAD would perform an edge-preserving filtering of the fMRI structural information, instead of enhancing its functional information. The RAD can be applied directly to SPM but, in this case, only a small improvement of the SPM quality can be achieved, because the originating fMRI is not taken into account. To overcome these difficulties, we propose to estimate the SPM from the noisy fMRI, compute the diffusion coefficients in the SPM space, and then perform the diffusion in the structural information-removed fMRI data using the coefficients previously computed. These steps are iterated until convergence. We have tested the new technique in both simulated and real fMRI images, yielding surprisingly sharp and noiseless SPMs with increased statistical significance. We also describe how to automatically estimate an appropriate scale parameter.

* Corresponding author. Fax: +55 11 3091 5718.
E-mail addresses: hae@lps.usp.br (H.Y. Kim), javier@lps.usp.br (J. Giacomantone), zcho@uci.edu (Z.H. Cho).

1077-3142/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2005.04.004

Keywords: Functional magnetic resonance imaging; fMRI; Anisotropic diffusion; Partial differential equation; Statistical parametric map; SPM

1. Introduction

The goal of functional neuroimaging is to map activities of a living brain in space and time. The gold standard for measuring brain cell activities involves direct and invasive electrical recording of membrane potential of individual neurons. However, such measurements are limited to certain experimental conditions. For studies on human subjects, non-invasive methods, such as positron emission tomography (PET) or functional magnetic resonance imaging (fMRI), have to be applied.

In the early 80s PET dominated the field of functional neuroanatomy, but in the past 10 years fMRI has developed into an alternative and powerful technique. Local increases in neural activity cause both a relative deoxygenation of blood and an increase in perfusion, that quickly reverses the deoxygenation, leading to an increase in oxygenation that endures for several seconds. fMRI is sensitive to the oxygenation of blood and has a spatio-temporal scale of about 1–3 mm and one or more seconds. The lower limits on the effective resolution of fMRI are physiological and imposed by the spatio-temporal organization of evoked hemodynamic responses (2–5 mm and 5–8 s).

Low signal-to-noise ratio in fMRI 4-D images compels us to use sophisticated image-processing techniques to detect activated brain areas. First, the data must pass through spatial transformations to correct subject's head-movement during fMRI acquisition. If the experiment involves different subjects, the data must also be normalized, that is, the images must "warp" such that they all conform to some standard brain.

After spatial transformations, statistical analyses are carried out. Many statistical procedures have been proposed to analyze blood oxygen level dependent fMRI data. Works [1,2] present a comparative study of these procedures. One of the most popular statistical procedures is the general linear model (GLM) [3]. In this model, the user specifies manually a design matrix and makes use of multiple linear regressions to estimate the parameters, i.e., to determine how well the time-series of each voxel fit the specified design matrix. These parameters are then used to compute the statistical significance of an effect. These statistics, spatially disposed, form a statistical parametric map (SPM). Course notes [4] is a good reference on fMRI processing.

Even with all these image-processing apparatuses, a noisy fMRI will yield a noisy SPM. To attenuate noise, classic linear or edge-preserving nonlinear image filters can be applied to either fMRI or SPM. Linear low-pass filters are routinely applied to fMRI data to attenuate noise. However, if fMRI is strongly low-pass-filtered, the resulting SPM will present blurred edges. Edge-preserving filters can be applied to SPM to filter noise [5], but they fail to substantially improve the quality of SPM because they do not take into account the originating fMRI data. If edge-preserving filters could be applied to fMRI, perhaps they would substantially improve the SPM. Unfortunately, if an edge-preserving filter is directly applied to 4-D fMRI,

it would enhance the fMRI structural information and erase the functional information, because the grayscale of an fMRI voxel depends primarily on the brain structure and only secondarily and faintly on the functional information. Thus, the goal of this paper is to find an edge-preserving filter applicable to fMRI to enhance the functional information.

In the literature, there are many works to attenuate noise and cluster activated regions in fMRI volumes [6–11]. In particular, Solé et al. [12] have recently proposed a technique named *anisotropic averaging*. This technique is inspired by the anisotropic diffusion [13]. Anisotropic averaging computes an initial set of clearly activated voxels. This set is then used to construct a complex similarity measure to compute the averaging coefficients.

In this paper, we present a new, simple, and elegant technique to obtain enhanced SPMs from noisy fMRI data that we called Robust Anisotropic Diffusion to compute enhanced Statistical Parametric Map (RADSPM). Instead of defining a highly complex similarity measure based on a set of clearly activated voxels, we use the gradient magnitudes estimated in the SPM space as the arguments of an edge-stopping function to compute the diffusion coefficients. The diffusion is then performed in the mean-corrected fMRI data (that is, the fMRI without the structural information), using the coefficients previously computed. A new SPM is then calculated and new diffusion coefficients are estimated. These steps are iterated a pre-defined number of times or until convergence.

The RADSPM, applied to both simulated and real fMRI data, has produced surprisingly sharp and noiseless SPMs. Our technique has substantially increased statistical significances of activated regions, which makes it possible to decide with more confidence if some brain region is activated or not. However, the use of the RADSPM seems not to be limited only to the binary activated/non-activated decisions, because the noiseless activation levels of SPMs produced by the RADSPM seems to bear more conclusive information than conventional noisy SPMs. The software used in this paper is publicly available for the research community [14].

2. Robust anisotropic diffusion

Scale-space is a theory used to process an image in multiple resolutions. Witkin [15] introduced a clean formalism for the linear scale-space. A coarse-scale image is obtained by low-pass filtering the fine scale image. Let $I(x,y)$: $\mathbb{R}^2 \to \mathbb{R}^+$ be an image in the continuous domain. The scale-space of this 2-D image is a 3-D image $I(x,y,t)$: $\mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{R}^+$ that satisfies the following partial differential equation:

$$\frac{\partial I(x,y,t)}{\partial t} = \operatorname{div}(\nabla I), \tag{1}$$

using the original image $I(x, y, 0) = I(x, y)$ as the initial condition. Variable t is an artificial time parameter that specifies the image scale. This linear scale-space has many nice mathematical properties. However, it blurs out image edges. To keep

edges sharp, while filtering noise and small details, Perona and Malik [13] defined the nonlinear anisotropic scale-space by modifying the partial differential Eq. (1) as follows:

$$\frac{\partial I(x,y,t)}{\partial t} = \operatorname{div}[g(\|\nabla I\|)\nabla I], \tag{2}$$

where $\|\nabla I\|$ is the gradient magnitude and g is an "edge-stopping" function. They suggested using one of the two edge-stopping functions below (all edge-stopping functions $g_i(x)$ presented in this paper have been dilated and scaled so that $g_i(0) = 1$ and their "influence functions" $xg_i(x)$ have local maximums at $x = 1$):

$$g_1(\|\nabla I\|) = \left[1 + \frac{\|\nabla I\|^2}{\sigma^2}\right]^{-1} \text{ or } g_2(\|\nabla I\|) = \exp\left[-\frac{\|\nabla I\|^2}{2\sigma^2}\right], \tag{3}$$

where σ is a positive constant. Perona and Malik discretized spatio-temporally their anisotropic diffusion Eq. (2) as:

$$I(s,t+1) = I(s,t) + \frac{\lambda}{|\eta_s|}\sum_{p\in\eta_s} g\big(|\nabla I_{s,p}(t)|\big)\nabla I_{s,p}(t), \tag{4}$$

where $I(s, t)$ is a discretely sampled image, s denotes the pixel position in a discrete 2-D or 3-D grid, t now denotes discrete iteration steps ($t \geqslant 0$), the constant $\lambda \in \mathbb{R}^+$ determines the rate of diffusion (usually $\lambda = 1$), and η_s represents the set of spatial neighbors of voxel s. For 2-D images, usually four neighboring voxels are considered: "north," "south," "west," and "east." For 3-D images, six voxels are usually considered: the above-mentioned four plus "up" and "down" voxels. Perona and Malik approximated the image gradient magnitude in a particular direction at iteration t as:

$$\nabla I_{s,p}(t) = I(p,t) - I(s,t), \quad p \in \eta_s. \tag{5}$$

The right choice of g can greatly affect the extent to which discontinuities are preserved. Black et al. [16] formulated the anisotropic diffusion as the problem of estimating a piecewise smooth image from noisy data and solved it using the robust statistics. Using robust estimation theory, they succeeded to define a better edge-stopping function (Tukey's biweight), and this is the edge-stopping function adopted in this paper:

$$g_3(\|\nabla I\|) = \begin{cases} \left[1 - \frac{\|\nabla I\|^2}{5\sigma^2}\right]^2, & \frac{\|\nabla I\|^2}{5} \leqslant \sigma, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

The resulting filtering process is called robust anisotropic diffusion (RAD). Using Tukey's function, the diffusion process converges faster and yields sharper edges than using Perona and Malik's. There have been some further developments in the anisotropic diffusion to better preserve edges (for example [17]), however, their aptness to filter fMRI remains subject of future research.

The scale of an image filtered by the anisotropic diffusion depends on the maximum number of iteration $t_{\max}$ and on the edge-stopping function scale parameter

σ. The appropriate selection of both parameters $t_{\max}$ and σ are still subjects of ongoing researches and no definitive solution seems to be available [18,19]. Using Perona–Malik's edge-stopping functions g_1 or g_2 the filtered image converges to an image with constant graylevel when $t_{\max} \to \infty$. On the other hand, the image filtered by RAD (using g_3) usually converges after a sufficient number of iteration steps, and consequently the scale of the filtered image depends practically only on σ, provided that the diffusion is iterated a suitable number of times. Some papers [19,16] present many different ways to choose an appropriate scale σ. In this paper, we will use the "robust scale" σ_e, suggested in [16]

$$\sigma_e = 1.4826 \;\; \mathrm{MAD}(\nabla I) = 1.4826 \;\; \mathrm{median}_I[|\|\nabla I\| - \mathrm{median}_I(\|\nabla I\|)|]. \tag{7}$$

The robust scale (as well as the scales chosen by any other automatic scale parameter setting techniques) must be considered only as an "initial guess" of the truly optimal scale. The optimal scale must be determined experimentally, starting from the robust scale, for a specific application. Our experiences (applying the RAD to 1-D signals, 2-D still images and fMRIs) have shown empirically that the optimal scale is usually two or three times larger than the robust scale.

3. General linear model

The general linear model (GLM) is a unified theoretical approach to perform classical statistical analysis of fMRI data [3]. Such models as one sample t test, two-sample t test, paired t test, Analysis of Variance (ANOVA), Analysis of Co-variance (ANCOVA), correlation [20], linear regression, multiple regression, and F test are particular cases of the GLM, depending on the adequate choice of predictor parameters. In this paper, we do not need all the generality of the GLM because we test the RADSPM only in single-subject block-design experiments. The GLM can be reduced in this case to a linear regression analysis and we describe in this section only this simplified case. However, the proposed technique can be extended straightforwardly to other cases. The GLM model can be stated as:

$$\begin{pmatrix} Y_1 \\ \cdot \\ Y_s \\ \cdot \\ Y_N \end{pmatrix} = \begin{pmatrix} X_{11} & \cdot & \cdot & X_{1L} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ X_{N1} & \cdot & \cdot & X_{NL} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \cdot \\ \beta_l \\ \cdot \\ \beta_L \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \cdot \\ \varepsilon_s \\ \cdot \\ \varepsilon_N \end{pmatrix}. \tag{8}$$

This equation is usually written in matrix notation as

$$Y = X\beta + \varepsilon. \tag{9}$$

In general, all variables involved in Eq. (9) are matrices. However, to simplify the exposition, we will suppose that Y is the column vector of an fMRI time-series of a fixed spatial position acquired at $t_1, \ldots, t_N$. X is the matrix of explanatory variables of the experiment, the stimulus or basis functions, called design matrix. This

matrix is defined after careful model specification according to the particular fMRI experiment and convolved with the hemodynamic response function to take into account the delay of the blood oxygenation effect. ε is the column vector of residual errors, which are considered identically distributed normal variables, independent or serially auto correlated. The amplitudes β_l of the basis functions are the parameters that we need to estimate.

The parameters β_l can be estimated by the least squares procedure. The following equation performs the least squares estimation of the parameters and also the maximum likelihood estimates

$$\hat{\beta} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}Y. \tag{10}$$

The parameter $\hat{\beta}$ can be transformed into the Student's t statistic τ by computing

$$\tau = \frac{c^{\mathrm{T}}\hat{\beta}}{\sqrt{\frac{(\varepsilon^{\mathrm{T}}\varepsilon)(X^{\mathrm{T}}X)^{-1}}{r}}}. \tag{11}$$

$c^{\mathrm{T}}\hat{\beta}$ is a linear combination of the estimated parameters and $r = N - p$ the degrees of freedom, where N is the number of volumes and p is the rank of the design matrix X. A statistical parametric map (SPM) is a 3-D image constructed by spatially disposing some particular kind of distribution values. Specifically, in this paper, we use Student's t statistical parametric map SPM$\{t\}$ obtained by spatially disposing τ.

The obtained statistic τ is used to perform a hypothesis test for each voxel in the 3-D image. Assuming that the null hypothesis H_0 indicates no correlation between the time-series Y and the design matrix X (H_0:$c^T\beta = 0$ and H_a:$c^T\beta > 0$ for the one-tail test), we would like to know how likely is our measure τ. The SPM allows us to perform a voxel by voxel comparison between the obtained value τ and the value τ_α corresponding to the selected significance level α (the acceptable false positive rate), accepting or rejecting the null hypothesis if $\tau < \tau_\alpha$ or $\tau \geqslant \tau_\alpha$, respectively. The following simple numerical example clarifies these ideas:

$$\begin{bmatrix} 50 \\ 51 \\ 60 \\ 62 \\ 51 \\ 52 \\ 62 \\ 63 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \end{bmatrix} \tag{12}$$

The first vector Y is the time-series of a voxel. Let us suppose that values were taken every 5 s. The second vector X is the design matrix indicating that we would like to detect a periodic activation that repeats every 20 s and lasts for 10 s. Estimating the parameter β with Eq. (10), we obtain $\hat{\beta}^{\mathrm{T}} = [10.75 \;\; 51]$. Applying Eq. (11), with

$c^{\mathrm{T}} = [1 \quad 0]$, we obtain the Student's t distribution with $r = N - p = 6$ degrees of freedom $\tau = 14.33$. This means that the voxel can be considered as activated, that is, correlated with the design matrix. The null hypothesis will be rejected at $\alpha = 0.01$ significance level $\tau_\alpha = 3.14$ for a one-tail t test.

4. Anisotropic averaging

The anisotropic diffusion is widely used in image processing to suppress noise and marginal edges in 2-D or 3-D grayscale images, while preserving important edges sharp. It can be applied directly to structural MRI [21], because a structural MRI can be regarded as a simple 3-D grayscale image. However, this technique cannot be applied directly to 4-D fMRI to obtain an enhanced SPM, because the direct application would perform an edge-preserving filtering of the fMRI structural information. The grayscale of an fMRI voxel depends primarily on the brain structure and only secondarily and faintly on the functional information. An edge-preserving filtering would enhance the structural information and regard the functional information as noise. The anisotropic diffusion can be applied directly to SPM [5], but it cannot improve substantially the quality of SPM because, in this case, the originating fMRI is not taken into account. Convolving fMRI data with a smoothing kernel generally increases the signal-to-noise ratio. However, this procedure also blurs the edges between activated and non-activated regions. Thus, it is desirable to perform only intra-region smoothing in fMRI data, avoiding inter-region smoothing.

Solé et al. [12] proposed this idea and called it anisotropic averaging, a technique motivated in part by the anisotropic diffusion. It consists in making a selective neighborhood averaging of the signal. Let I be an fMRI image and let $I(s, n)$ denote the voxel value at spatial voxel position s and volume (i.e., scan, observation) n. The anisotropic averaging will gradually modify I. Let us denote the fMRI image at iteration $t \geqslant 0$ as $I(s, n, t)$, with $I(s, n, 0) = I(s, n)$. Then, the selective neighborhood averaging consists in:

$$I(s, n, t+1) = \frac{1}{\sum_{p \in \eta_s} w(s, p)} \sum_{p \in \eta_s} w(s, p) I(p, n, t), \tag{13}$$

for all voxels s, all volumes n and $t \geqslant 0$. As before, η_s represents the set of spatial neighbors of voxel s.

Let $I(s)$ denote the time-series signal at voxel s. A similarity measure Ψ determines the averaging weights $w(s, p)$ using the time-series of the signal being averaged $I(s)$ and its neighbor $I(p)$

$$w(s, p) = \Psi(I(s), I(p)). \tag{14}$$

This similarity measure allows us to distinguish activated voxels from non-activated ones. It allows us to perform a selective averaging, combining only signals of the same class. Solé et al. propose to compute the initial set Ω of clearly activated voxels by thresholding the correlation coefficients. Then, the Fourier spectra of voxels in Ω are computed to define the similarity measure function Ψ. The

Fourier spectrum of time-series of each voxel s is also computed to evaluate the similarity between s and the voxels in Ω. The whole procedure is highly complex and the readers are referred to [12] for further details. After each step of the neighborhood averaging (13), a new set Ω of clearly activated voxels is computed.

5. The proposed method

We propose a different approach, directly related to the RAD, that we named RADSPM. Our method is simple, elegant, and has yielded surprisingly enhanced SPMs when applied to both simulated and real fMRI data. The RADSPM has also increased substantially the statistical significance of activated regions, which makes it possible to decide with more confidence if a certain brain region is activated or not.

Let be given an fMRI I' (that contains both structural and activation information) and let $I'(s, n)$ be the voxel value of I' at spatial voxel position s and volume n. First of all, the brain structural information is removed from I', yielding the mean-corrected fMRI I

$$I(s,n) = I'(s,n) - \text{average}_n(I'(s,n)), \quad \text{for all } s \text{ and } n. \tag{15}$$

This pre-processing is important, because structural and functional regions of the brain do not necessarily match. Only the activation information should be diffused between intra-region voxels, avoiding inter-region and structural information diffusions. If the structural information is diffused together with the activation information, a blurred SPM will result. Note that the activation information is not affected at all by the mean-correction.

Let us denote the fMRI data at iteration $t \geqslant 0$ of the diffusion process as $I(s, n, t)$, where $I(s, n, 0)$ is the initial mean-corrected and normalized fMRI. The RADSPM algorithm is described below:

1. Let $t \leftarrow 0$.
2. Using the fMRI $I(s, n, t)$ and the design matrix X, estimate the Student's t statistic τ for every voxel s, obtaining the SPM T (Eqs. (10) and (11)). Let us denote the value of the SPM T at voxel s and iteration t as $T(s, t)$.
3. Compute the diffusion coefficients. The diffusion coefficient between a voxel s and its neighboring voxel p at instant t is

$$g\big(|\nabla T_{s,p}(t)|\big), \quad \text{where } \nabla T_{s,p}(t) = T(p,t) - T(s,t). \tag{16}$$

4. Use these coefficients to perform the diffusion in $I(s, n, t)$, yielding the diffused fMRI $I(s, n, t + 1)$ at iteration $t + 1$

$$I(s,n,t+1) \leftarrow I(s,n,t) + \frac{\lambda}{|\eta_s|} \sum_{p \in \eta_s} g\big(|\nabla T_{s,p}(t)|\big) \nabla I_{s,p}(n,t), \tag{17}$$

where $\nabla I_{s,p}(n, t) = I(p, n, t) - I(s, n, t)$.

5. Let $t \leftarrow t + 1$ and repeat steps 2–6 until the average of diffused values (the second term of Eq. (17)) is below some predefined threshold. It is also possible to specify the desired number of iteration steps $t_{\max}$, instead of defining the minimal average diffused value.

In our preliminary paper [22], we suggested computing the diffusion coefficients in the space of estimated parameters $\hat{\beta}$ (instead of computing them in the SPM space). Indeed, computing diffusion coefficients in the space of parameters yields a better-quality SPM for simulated fMRIs, where the noise and the activation are completely unrelated. However, for real fMRI data, experimental results have shown that it is better to compute the diffusion coefficients in the SPM space.

The processing time is the weak point for the practical use of the RADSPM. As we describe in next section, tens of RADSPM iteration steps are necessary to obtain a high-quality SPM. As it is necessary to compute a new SPM in each repetition, the proposed method is many tens of times slower than the conventional method. Consequently, the improvement of our algorithm's performance is subject of future research, even though we consider that, as computers become faster and faster, the processing time may not be a critical aspect in not a distant future.

6. Experimental results

6.1. Simulated fMRI

In the first experience, we generated a simple artificial 4-D fMRI with $3 \times 10 \times 10$ voxels per volume and 84 volumes, where blocks of 6 non-activated volumes alternated with blocks of 6 activated ones, starting with rest. Voxel values were 16,000 added to Gaussian noise with zero mean and standard deviation 4000. Activated volumes had an activated $3 \times 6 \times 6$ region in the center, with two non-activated $3 \times 2 \times 2$ subregions. Activated voxels had their values increased by 5000. The signal-to-noise ratio between the mean-corrected noiseless phantom I and the mean-corrected noisy phantom $\hat{I}$, computed according to the formula below, was found to be $-9.5\,\mathrm{dB}$:

$$\mathrm{SNR} = -10 \log_{10}\left[\frac{\sum_n \sum_s (I(s,n) - \hat{I}(s,n))^2}{\sum_n \sum_s (I(s,n))^2}\right] \mathrm{dB}. \tag{18}$$

This simple phantom may not simulate perfectly all the noise effects present in a real fMRI. However, it can be used for testing purposes of the proposed algorithm.

First column of Fig. 1 depicts the original SPM obtained by the classic algorithm. The robust scale of this SPM is $\sigma_e = 1.092$. Second column depicts the SPM generated by the RADSPM, using an adequate scale parameter ($\sigma = 3 = 2.747\sigma_e$, $t = 90$). Third column depicts the SPM generated by the RADSPM, using too large a scale parameter ($\sigma = 5 = 4.579\sigma_e$, $t = 90$). Statistics with $0 \leqslant \tau \leqslant 20$ were linearly mapped into the grayscale ranging from 0 to 255 (voxels with $\tau \leqslant 0$ are depicted as black and

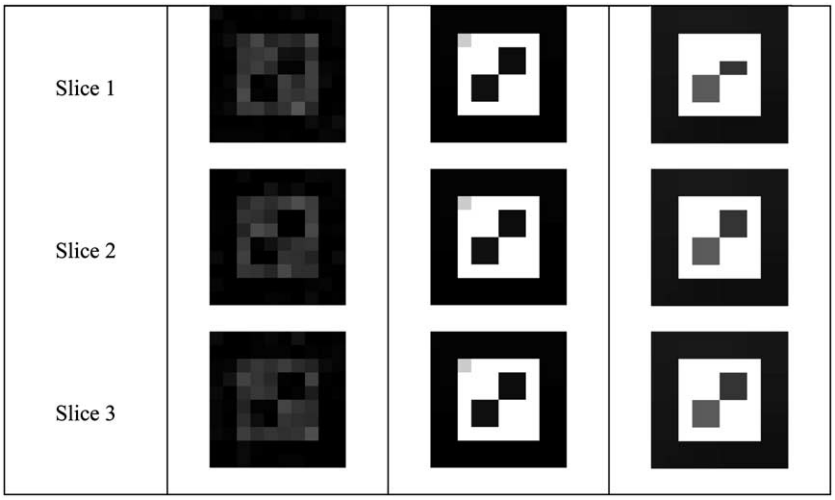

Fig. 1. SPMs obtained from the simulated fMRI. First column: the SPM obtained by the conventional algorithm, with the robust scale $\sigma_e = 1.092$. Second column: the SPM generated by the RADSPM using an adequate parameter ($\sigma = 3 = 2.747\sigma_e$, $t = 90$). Third column: the SPM generated by the RADSPM using too large a parameter ($\sigma = 5 = 4.579\sigma_e$, $t = 90$). Statistics with $0 \leqslant \tau \leqslant 20$ were linearly mapped into grayscale from 0 to 255 (voxels with $\tau \leqslant 0$ are depicted as black and $\tau \geqslant 20$ are depicted as white).

$\tau \geqslant 20$ are depicted as white). Note that the filtered SPM is almost completely noiseless, and the edges are perfectly preserved, provided that an adequate scale parameter is chosen.

Table 1 shows the maximum, the average and the minimum values of activated and non-activated regions of SPMs obtained using different parameters σ (the number of iteration steps was fixed at $t = 90$). Consider, for example, $\sigma = 3$. The smallest value τ of activated voxels has increased from the original 3.37 to 16.73. Meanwhile,

Table 1
Information about the SPMs obtained from the simulated fMRI using different scale parameters σ (with $t = 90$ fixed)

	Original	1.0	2.0	3.0	4.0	5.0	6.0
Activated							
Maximum	8.49	25.46	31.32	43.93	49.24	49.16	44.31
Average	5.67	15.30	22.73	33.74	44.01	46.05	43.95
Minimum	3.37	3.71	5.52	16.73	39.72	42.68	43.60
Non-activated							
Maximum	2.57	4.41	1.29	2.17	4.25	46.96	44.17
Average	0.00	−0.73	−0.37	0.17	1.46	3.44	8.54
Minimum	−2.92	−4.41	−1.50	−0.99	0.34	1.93	3.67

The column "original" refers to the SPM obtained by the traditional method. Other columns refer to the SPMs obtained by the RADSPM.

Table 2
Information about the SPMs obtained from the simulated fMRI using different number of iteration steps t (with $\sigma = 3$ fixed)

	Original	1	4	40	90	160	320
Activated							
Maximum	8.49	15.29	30.05	40.92	43.93	47.53	48.77
Average	5.67	11.91	20.66	31.53	33.74	35.56	39.89
Minimum	3.37	8.16	14.48	16.73	16.73	16.73	16.73
Non-activated							
Maximum	2.57	5.59	6.32	2.17	2.17	2.17	2.17
Average	0.00	0.15	0.16	0.14	0.17	0.20	0.22
Minimum	−2.92	−3.18	−3.20	−1.44	−0.99	−0.54	−0.14

After sufficient steps (say, $t = 40$), the output practically remains constant.

the largest τ of non-activated voxels has decreased a little. This means that one can decide whether a voxel is activated or not with much more confidence using the SPM generated by our method. The average value of activated voxels has increased from 5.67 to 33.74, while the average value of non-activated voxels has remained almost constant, meaning again that our method has improved the quality of the SPM. The quality begins to deteriorate at $\sigma = 4$ and gets utterly worse at $\sigma = 5$: the highest τ of non-activated voxels has increased to 4.25 and 46.96, respectively, indicating that the edges between activated and non-activated regions began to melt. As the best quality was obtained at $\sigma = 3 = 2.747\sigma_e$, it seems that the optimal scale is situated somewhere between $2\sigma_e$ and $3\sigma_e$.

Table 2 shows that, after sufficient iteration steps (say, $t = 40$), the output of the RADSPM practically does not change with more steps. We tested this fact using scale parameter $\sigma = 3$.

6.2. Real fMRI #1

Fig. 2 depicts the SPMs computed from a real fMRI and overlaid on the structural MRI. The SPMs were generated using the conventional and RADSPM techniques.

Whole-brain Blood Oxygenation Level Dependent/Echo-Planar Imaging (BOLD/EPI) fMRI data were acquired on a 1.5 T Philips Eclipse system. Each acquisition consisted of 25 contiguous slices ($64 \times 64 \times 25$, 4 mm × 4 mm × 5 mm voxels). Acquisition took 180 s, with the scan-to-scan repeat time (TR) set to 3 s and echo time (TE) set to 35 ms. Sixty acquisitions were made, in blocks of 6, giving 10 acquisitions (30 s) per block. The condition for successive blocks alternated between rest and visual stimulation, starting with rest. Visual stimulation was given with flickering red LED light with a frequency of 2 Hz. First five acquisition data were discarded due to T1-effects induced instability. The volumes were realigned (i.e., motion-corrected) and normalized using SPM99 program [23]. Sinc interpolation was used in both realignment and normalization. The resulting fMRI data had $79 \times 95 \times 68$ voxels per volume and 55 volumes. Although this is a single-subject study, we performed

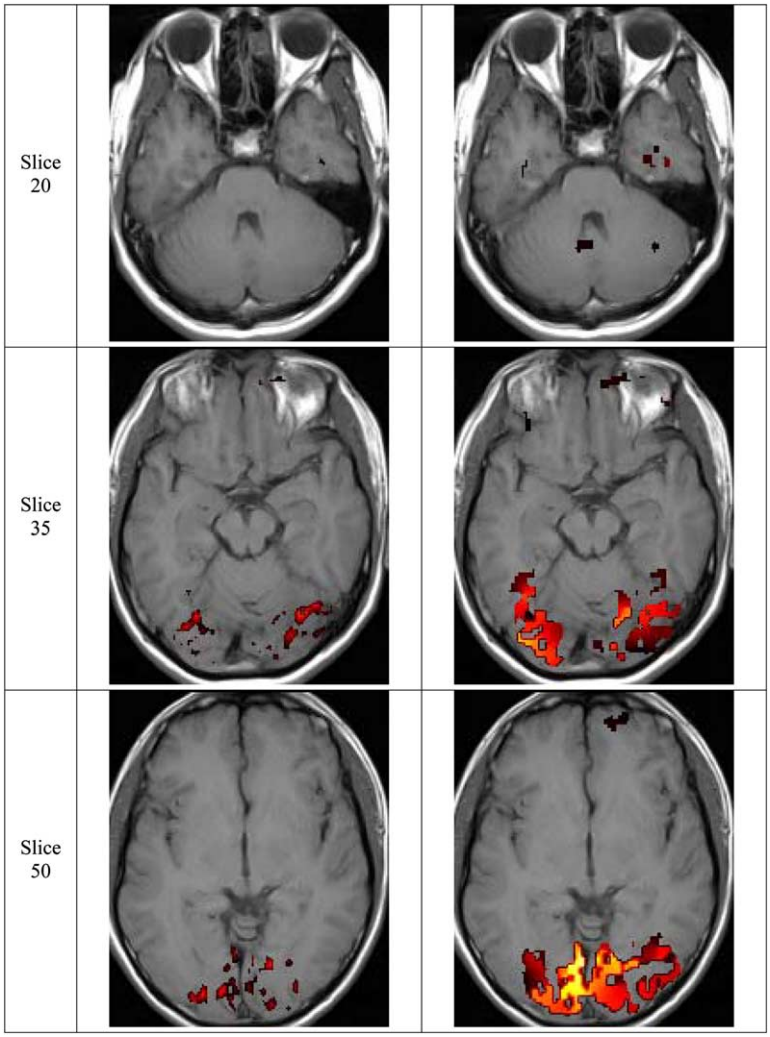

Fig. 2. SPMs computed from the real visual-stimulation fMRI #1 and overlaid on the structural MRI. Voxels with $6 \leqslant \tau \leqslant 20$ were mapped into color scale called "hot." Left column: the SPM generated by the conventional technique, with the robust scale $\sigma_e = 0.629$. Right column: the SPM generated by the RADSPM using adequate parameters ($\sigma = 1.5 = 2.39\sigma_e$, $t = 90$).

the normalization to overlay the functional SPMs over the structural MRI. We also obtained a spin echo T1-weighted image to be used as the structural image. It was acquired in 256×256 matrix with 26 slices. Each slice has a 5 mm thickness and 0.78 mm $\times$ 0.78 mm in-plane resolution. After the normalization, the resolution of the structural MRI was $157 \times 189 \times 136$ voxels.

The design matrix was constructed according to the acquisition conditions and convolved with the hemodynamic response function. Then, the SPMs were generated

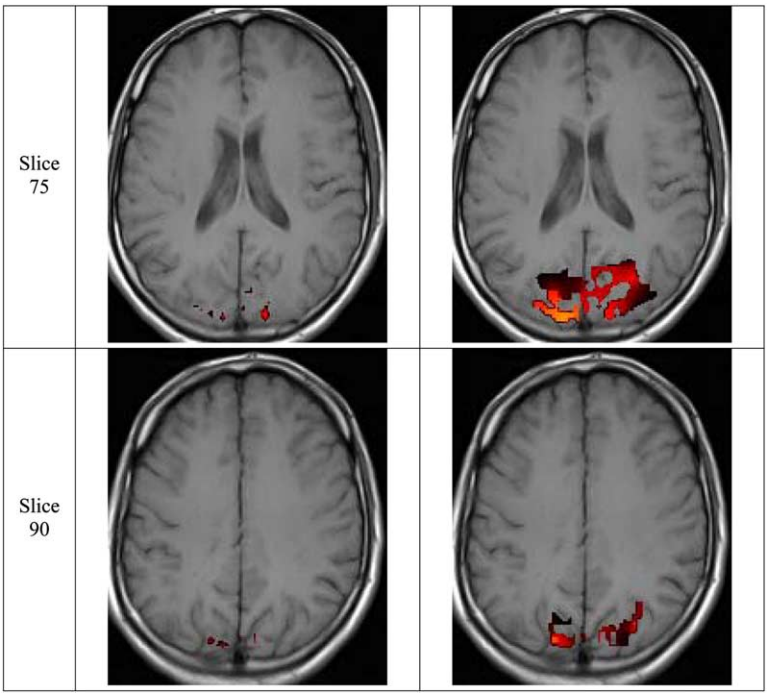

Fig 2. (*continued*)

using the conventional and RADSPM techniques. To certify that our program had no implementation errors, the conventional SPM generated by our program was compared with the SPM generated by SPM99 program [23] and found to be virtually identical. Then, the SPMs were overlaid on the structural MRI. We used the program MRIcro [24] to overlay the functional SPMs over the structural MRI and generate slice images. We adopted a positive activation range with the initial value $\tau = 6$ and the final value $\tau = 20$ (that is, voxels with $6 \leqslant \tau \leqslant 20$ were mapped into color scale called "hot"). Left column of Fig. 2 depicts some slices of the conventional SPM overlaid on the structural MRI. The robust scale of this SPM was $\sigma_e = 0.629$. Right column of Fig. 2 depicts the SPM generated by the RADSPM using adequate parameters ($\sigma = 1.5 = 2.39\sigma_e$, $t = 90$).

In a real fMRI, we do not know the ideal SPM, neither which voxels are actually activated. Consequently, it is difficult to state that an SPM is better than another. However, visually, it seems that the proposed technique has improved substantially the quality of the SPM. We cannot obtain the same kind of information that we presented for the simulated fMRI in Tables 1 and 2. However, analogous information could be obtained. Table 3 presents the maximum, the average, and the minimum values of clearly activated voxels: the 3% highest-valued voxels. It also presents the maximum, the average, and the minimum absolute values of clearly non-activated voxels: the 30% voxels with the lowest absolute values (we use absolute values to

Table 3
Information about clearly activated region (the 3% highest-valued voxels) and clearly non-activated region (the 30% voxels with the lowest absolute values) of SPMs obtained from the real fMRI #1 using different scale parameters σ (with $t = 90$ fixed)

	Original	0.3	0.7	1.1	1.5	1.7	1.9
Highest 3%							
Maximum	15.22	16.08	21.06	21.79	21.85	23.23	22.78
Average	5.72	7.25	10.14	11.78	12.38	12.37	12.15
Minimum	3.82	5.38	7.79	8.77	9.66	9.43	8.71
Absolute value of the lowest 30%							
Maximum	0.52	1.13	1.52	0.98	0.69	0.68	0.69
Average	0.26	0.57	0.78	0.50	0.35	0.34	0.35
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Column "original" refers to the SPM generated by the conventional method and other columns refer to the SPMs generated by the RADSPM.

not mix negatively activated and non-activated voxels). This information is presented for the original SPM and the SPMs generated by the RADSPM, using different parameters σ (the number of iteration steps was fixed at $t = 90$). Consider, for example, $\sigma = 1.5$. The smallest τ value of clearly activated voxels has increased from the original 3.82 to 9.66. Meanwhile, the largest τ of clearly non-activated voxel has remained nearly constant (0.52–0.69). The average value of clearly activated voxels has increased from 5.72 to 12.38, while the average value of clearly non-activated voxels remained almost constant (0.26–0.35). This seems to indicate that the RADSPM detects the clearly activated voxels with more statistical confidence, without damaging the detection of the clearly non-activated voxels. The lowest and the average values of clearly activated voxels are maximal at $\sigma = 1.5 = 2.39\sigma_e$, indicating that the optimal scale parameter may be situated somewhere between $2\sigma_e$ and $3\sigma_e$.

Table 4 shows that, after sufficient iteration steps (say, $t = 90$), more steps practically do not modify the output of the RADSPM. The scale parameter was fixed at $\sigma = 1.5$.

Table 4
Information about the SPMs generated from the real fMRI #1 using different number of iteration steps t (with $\sigma = 1.5$ fixed)

	Original	1	4	30	90	130	170
Highest 3%							
Maximum	15.22	16.05	16.89	21.78	21.85	23.70	22.47
Average	5.72	6.32	7.49	10.68	12.38	12.77	12.98
Minimum	3.82	4.28	5.18	7.79	9.66	10.02	10.25
Absolute value of the lowest 30%							
Maximum	0.52	0.56	0.64	0.75	0.69	0.68	0.67
Average	0.26	0.28	0.31	0.37	0.35	0.34	0.34
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00

After sufficient steps (say, $t = 90$), the output practically remains constant.

6.3. Real fMRI #2

Fig. 3 depicts the SPMs computed from another real fMRI, acquired, and processed similarly but independently from the real fMRI #1. Once more, performing a visual inspection, it seems that the proposed technique has substantially improved the quality of the SPM. The robust scale of the original SPM is $\sigma_e = 0.433$. We used scale parameter $\sigma = 1.1 = 2.54\sigma_e$ to obtain an adequately filtered SPM.

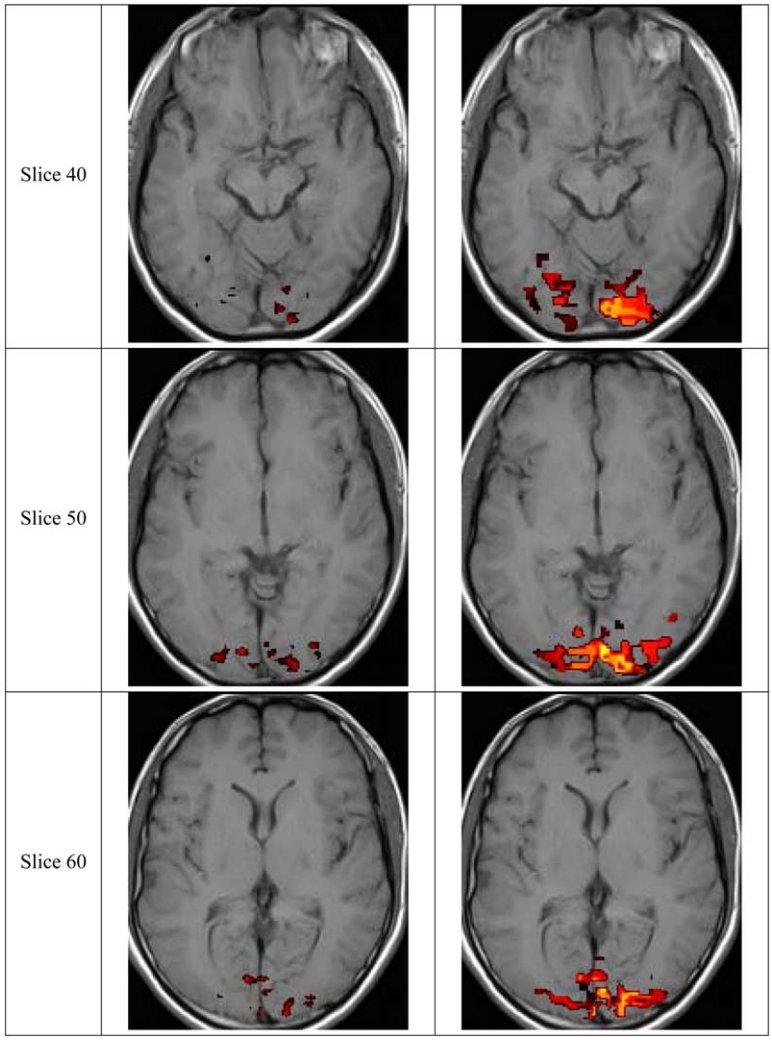

Fig. 3. SPMs computed from the real visual-stimulation fMRI #2 and overlaid on the structural MRI. Voxels with $6 \leqslant \tau \leqslant 20$ were mapped into color scale called "hot." Left column: the SPM generated by the conventional technique, with the robust scale $\sigma_e = 0.433$. Right column: the SPM generated by the RADSPM using adequate parameters ($\sigma = 1.1 = 2.54\sigma_e$, $t = 90$).

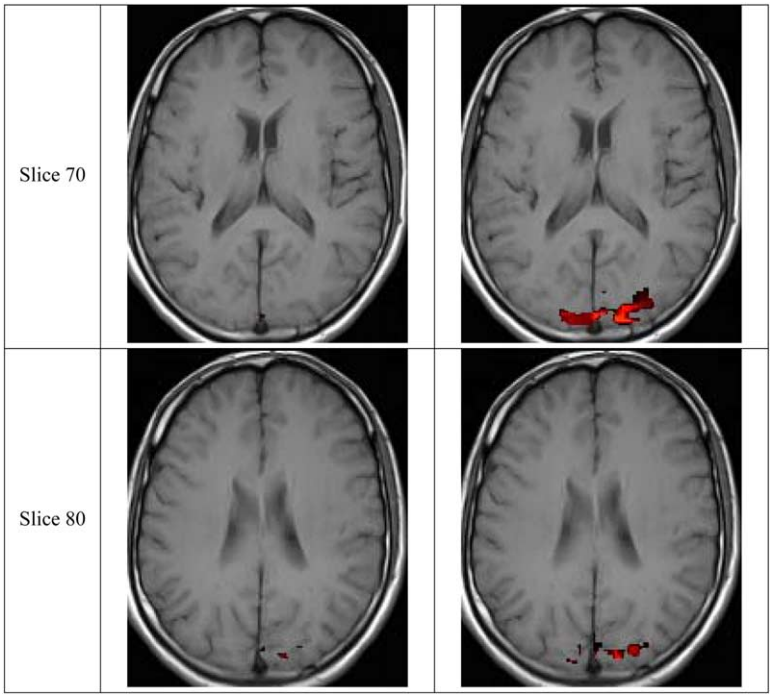

Fig 3. (*continued*)

Table 5 presents the maximum, the average, and the minimum values of clearly activated and clearly non-activated regions. The lowest value of the clearly activated region is maximal at $\sigma = 0.7 = 1.62\sigma_e$ and the average value of the same region is maximal at $\sigma = 1.5 = 3.46\sigma_e$, indicating once again that the optimal scale parameter may be situated somewhere between $2\sigma_e$ and $3\sigma_e$.

Table 5
Information about the SPMs obtained from the real fMRI #2 using different scale parameters σ (with $t = 90$ fixed)

	Original	0.3	0.7	1.1	1.5	1.7	1.9
Highest 3%							
Maximum	16.21	16.21	16.48	18.85	18.06	16.42	15.72
Average	4.51	5.80	7.78	8.72	9.05	8.96	8.77
Minimum	2.65	3.95	4.86	4.85	4.22	3.86	3.81
Absolute value of the lowest 30%							
Maximum	0.48	1.00	1.14	1.02	1.12	1.12	1.14
Average	0.24	0.52	0.58	0.54	0.58	0.58	0.59
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Column "original" refers to the SPM generated by the conventional method and other columns refer to the SPMs generated by the RADSPM.

Table 6
Information about the SPMs generated from the real fMRI #2 using different number of iteration steps t (with $\sigma = 1.1$ fixed)

	Original	1	4	30	90	130	170
Highest 3%							
Maximum	16.21	16.21	16.17	16.42	18.85	18.74	18.37
Average	4.51	4.88	5.61	7.73	8.72	9.01	9.16
Minimum	2.65	2.88	3.34	4.40	4.85	5.01	5.01
Absolute value of the lowest 30%							
Maximum	0.48	0.52	0.59	0.74	1.02	1.23	1.39
Average	0.24	0.25	0.29	0.36	0.54	0.66	0.79
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00

After sufficient steps (say, $t = 90$), the output practically remains constant.

Table 6 shows that, after sufficient iteration steps (say, $t = 90$), the output of the RADSPM virtually does not change with more steps. The scale parameter was fixed at $\sigma = 1.1$.

7. Conclusion

In this paper, we have presented a new, simple, and elegant technique named RADSPM to obtain enhanced SPMs from noisy fMRI. It is directly inspired by the robust anisotropic diffusion. Experimental results have shown that the RADSPM generates surprisingly sharp and noiseless SPMs. The proposed technique has also increased the statistical significances of the activated regions, what makes it possible to decide with more confidence if some regions of the brain are activated or not. The RADSPM is sensitive to the scale parameter selection. We concluded experimentally that the optimal scale parameter seems to be situated somewhere between $2\sigma_e$ and $3\sigma_e$, where σ_e is the robust scale of the conventional SPM. We also concluded that the choice of the number of iteration steps t is not critical, provided that enough steps are executed.

Acknowledgments

We express our gratitude to Young Don Son and Chang Ki Kang for their help and support in acquiring the fMRI data used in this paper. H.Y. Kim and J.O. Giacomantone thanks FAPESP and CNPq for the partial financial supports of this work under Grants 2003/13752-9, 03/03369-3, and 305065/2003-3.

References

[1] N. Lange, S.C. Strother, J.R. Anderson, F.A. Nielsen, A.P. Holmes, T. Kolenda, R. Savoy, L.K. Hansen, Plurality and resemblance in fMRI data analysis, NeuroImage 10 (1999) 282–303.

[2] S. Gold, B. Christian, S. Arndt, G. Zeien, T. Cizadlo, D.L. Johnson, M. Flaum, N.C. Andreasen, Functional MRI statistical software packages: a comparative analysis, Hum. Brain Mapp. 6 (1998) 73–84.
[3] K.J. Friston, A.P. Holmes, K.J. Worsley, J.P. Poline, C.D. Frith, R.S.J. Frackowiak, Statistical parametric maps in functional imaging: a general linear approach, Hum. Brain Mapp. vol. 2 (1995) 189–210.
[4] K.J. Friston (Ed.), SPM Course—Short Course Notes. Available at <http://www.fil.ion.ucl.ac.uk/spm/course/>
[5] H.S. Neoh, G. Sapiro, Using Anisotropic Diffusion of Probability Maps for Activity Detection in Block-Design Functional MRI, in: Proceedings of the IEEE International Conference on Image Processing I, vol. 1, 2000, pp. 621–624
[6] C. Goutte, P. Toft, E. Rostrup, F.A. Nielsen, L.K. Hansen, On clustering fMRI time series, NeuroImage 9 (3) (1999) 298–310.
[7] B.A. Ardekani, I. Kanno, Statistical methods for detecting activated regions in functional MRI of the brain, Magn. Reson. Imaging 16 (10) (1998) 1217–1225.
[8] J. Kershaw, B.A. Ardekani, I. Kanno, Application of Bayesian inference to fMRI data analysis, IEEE Trans. Med. Imaging 18 (1999) 1138–1153.
[9] K.H. Chuang, M.J. Chiu, C.C. Lin, J.H. Chen, Model-free functional MRI analysis using kohonen clustering neural network and fuzzy c-means, IEEE Trans. Med. Imaging 18 (1999) 1117–1128.
[10] K.J. Friston, P. Jezzard, R. Turner, The analysis of functional MRI time-series, Hum. Brain Mapp. 1 (1994) 153–171.
[11] J. Baudewig, P. Dechent, K.D. Merboldt, J. Frahm, Thresholding in correlation analyses of magnetic resonance functional neuroimaging, Magn. Reson. Imaging 21 (2003) 1121–1130.
[12] A.F. Solé, S.C. Ngan, G. Sapiro, X.P. Hu, A. López, Anisotropic 2-D and 3-D averaging of fMRI signals, IEEE Trans. Med. Imaging 20 (2) (2001) 86–93.
[13] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE. Trans. Pattern Anal. Mach. Intell. 12 (7) (1990) 629–639.
[14] An implementation of the RADSPM, Available at <http://www.lps.usp.br/~hae/software/fmri>
[15] P. Witkin, Scale-Space Filtering, in: Proceedings of 8th International Joint Conference Artificial Intelligence , vol. 2, 1983, pp. 1019–1022
[16] M.J. Black, G. Sapiro, D.H. Marimont, D. Hegger, Robust anisotropic diffusion, IEEE Trans. Image Process. 7 (3) (1998) 421–432.
[17] P.K. Saha, J.K. Udupa, Scale-based diffusive image filtering preserving boundary sharpness and fine structures, IEEE Trans. Med. Imaging 20 (11) (2001) 1140–1155.
[18] V. Solo, A Fast Automatic Stopping Criterion for Anisotropic Diffusion, in: Proceedings of ICASSP, IEEE, vol. 2, 2002, pp. 1661–1664
[19] F. Voci, S. Eiho, N. Sugimoto, H. Sekiguchi, Estimating the gradient threshold in the Perona–Malik equation, IEEE Signal Processing Mag. (2004) 39–46.
[20] P.A. Bandettini, A. Jesmanowick, E.C. Wong, J.S. Hyde, Processing strategies for time-course data sets in functional MRI of the human brain, Magn. Reson. Med. 30 (1993) 161–173.
[21] G. Gerig, O. Kübler, R. Kikinis, F.A Jolesz, Nonlinear anisotropic filtering of MRI data, IEEE Trans. Med. Imaging 11 (2) (1992) 221–232.
[22] H.Y. Kim, Z.H. Cho, Robust Anisotropic Diffusion to Produce Clear Statistical Parametric Map from Noisy fMRI, in: Proceedings of Sibgrapi—Brazilian Symposium on Computer Graphics and Image Processing, 2002, pp. 11–17
[23] Program SPM'99, University College London, Available at <http://www.fil.ion.ucl.ac.uk/spm/software/spm99/>
[24] MRIcro version 1.37, University of South Carolina, Available at <http://www.mricro.com>