Journal of Mathematics and Statistics 4 (4): 255-263, 2008
ISSN 1549-3644

Application of MultiScale Hidden Markov Modeling Wavelet Coefficients to fMRI Activation Detection

[1]Fangyuan Nan, [1]Yaonan Wang and [2]Xiaoping Ma
[1]College of Electrical and Information Engineering, Hunan University, Changsha, China
[2]School of Information and Electrical Engineering, China University of Mining and Technology, China

Abstract: Problem Statement: The problem of detection of functional magnetic resonance images (fMRIs), that is, to decide active and nonactive regions of human brain from fMRIs is studied in this paper. fMRI research is finding and will find more and more applications in diagnosing and treating brain diseases like depression and schizophrenia. At its initial stage fMRI detection are pixel-wise methods, which do not take advantage of mutual information among neighboring pixels. Ignoring such spatial information can reduce detection accuracy. During past decade, many efforts have been focusing on taking advantage of spatial correlation inherent in fMRI data. Most well known is smoothing using a fixed Gaussian filter and the compensation for multiple testing using Gaussian random field theory as used by Statistical Parametric Mapping (SPM). Other methods including wavelets had also been proposed by the community. **Approach:** In this study a novel two-step approach was put forward that incorporates spatial correlation information and is amenable to analysis and optimization. First, a new multi scale image segmentation algorithm was proposed to decompose the correlation image into several different regions, each of which is of homogeneous statistical behavior. Second, each region will be classified independently as active or inactive using existing pixel-wise test methods. The image segmentation consists of two procedures: Edge detection followed by label estimation. To deduce the presence or absence of an edge from continuous data, two fundamental assumption of our algorithm are 1) each wavelet coefficient was described by a 2-state Gaussian Mixture Model (GMM); 2) across scale, each state is caused by its parent state, hence the Multiscale Hidden Markov Model (MHMM). The states of Markov chain are unknown ("hidden") and represent the presence (state 1) or absence (state 0) of edges. Using this interpretation, the edge detection problem boils down to the posterior state estimation given observation. **Results:** Data processing results demonstrate much improved efficiency of our algorithm compared with pixel-wise detection methods. **Conclusions:** Our methods and results stress the importance of spatial-temporal modeling in fMRI research.

Key words: Functional magnetic resonance imaging (fMRI), wavelet analysis, image segmentation, hidden Markov model, spatial-temporal modeling

INTRODUCTION

Spatial modeling of fMRI and outline fort the method: Magnetic Resonance Imaging (MRI) is a powerful diagnostic imaging technique based on the principle of nuclear magnetic resonance, describing the interaction of nuclei and magnetic fields. While traditional MRI provides only static images to analyze anatomical structure, functional MRI (fMRI), a newer imaging modality which is based on MRI and just came to the stage around two decades ago, acquires a series of images to detect neural activity, to locate brain activation. In other words, the central task for fMRI is to obtain maps of active and nonactive regions of the brain[1-2].

Closely related to this study is[1], which uses generalized likelihood ratio test to get the activation map. Since its publication[1], has received many citations, including, but not limited to[26,28,29,30]. It almost forms the basis for fMRI complex data processing. In particular, the model and methods in[1] were also validated and generalized by some other work[29,30].

As in[1], pixelwise detection for fMRI is most common in practice[2-4]. Some standards software's like SPM and AFNI[13-14] now already exist. However, these techniques do not take advantage of mutual information between neighboring pixels. Ignoring such spatial information can cause problems. For example, at first sight to a physician, the last figure in[1] seems surprising

Corresponding Author: Fangyuan Nan, College of Electrical Information Engineering, Hunan University, Changsha, China

because there are activation areas outside brain! On the other hand, utilizing spatial information may enhance our detection accuracy. For example, it may be quite possible that an activated (contiguous) areas is larger than individual pixel dimensions. In other words, activated areas tend to occur in clusters of neighboring pixels, or if we know, by some means, there is strong indication that a large group of pixels, which may be thought of as one large pixel at very coarse (spatial) scale, is active, then the individual pixels inside this group, which may be regarded as pixels at a finer scale, are more likely to be active themselves. Hence comes the idea of (spatial) scale and incorporating spatial correlation into the fMRI detection process.

In light of the above simple idea, the kind of the pixelwise detection is oversimplistic. Therefore it is necessary to develop detection methods taking advantage of spatial correlation. There are many approaches to attack the problem, for example, cluster analysis[23] and Independent Component Analysis (ICA) algorithm[6]. Detection methods using Bayesian strategies have been proposed for fMRI[22]. Just as in pixelwise detection, we need to model each time series; when we turn attention to spatial correlation, we also need consider spatial modeling for our problem. This is by no means an easy job. But, we note that those Bayesian methods mentioned above are all restricted to modeling on the finest scale. Such methods tend to be very computationally demanding and are often difficult to analyze and interpret. Therefore, multiscale modeling (specifically, multiscale image segmentation) will be considered that incorporates spatial correlation information and is much more amenable to analysis and optimization.

Some work has already been done in this aspect[10-11, 28]. This study will adopt a two-step approach for fMRI detection: multi-scale image segmentation will be first used to break the correlation image into several different regions, each of which is of homogenous statistical behavior, then these regions will be classified independently as active or inactive by single-pixel detection methods. Since pixelwise detection has been elaborated in other literature, this study will concentrate on the first step of image segmentation.

MATERIAL AND METHODS

General setting of Bayesian image segmentation: In essence, the detection of fMRI, in a more general setting, may be addressed as particular image segmentation problem.

We are given a random continuous noisy image Y which must be segmented into a discrete image X consisting of regions of distinct statistical behavior. For example, in fMRI, image Y may be composed of the correlation data (specifically the correlation between amplitude time series and the reference signal) at all pixels, image X may be just the binary detection map.

We assume that each observed pixel in image Y is dependent on a corresponding unobserved label in X. Each label specifies one of M possible states, each with its own statistical behavior. In our case, ideally we hope M = 2 indicating "active" and "inactive". However, in the general case of our two-step approach, in the first step, we just want to segment the correlation image in to regions with homogeneous statistical behavior, so M is not confined to be 2. The decision on activeness is made in the second step, i.e., applying single-pixel detection to each region.

The dependence of observed pixels on their labels is specified through the conditional distribution of Y given X, i.e., $P_{y|x}(y|x)$.This contains most information, but is very difficult to determine in practice. A Priori knowledge about the size and shapes of regions will be modeled by the a priori distribution p(x). The complexity of the problem is easily imagined because even this information eludes us for real problems. Under some criterion (for example, maximize A Posteriori (MAP), we wan to find the estimate of X given observed image Y = y. This is the general framework for Bayesian image segmentation problem.

By Bayes formula, the estimate of X given observed image Y = y is:

$$\hat{x} = \arg\max_x pX\left|Y(x,y) = \max_x \frac{px(x)py\left|x(y,x)\right.}{py(y)} \right. \quad (1)$$

$$= \max_x \frac{p(X=x)p(Y=y\left|X=x\right.)}{\sum_{all\,X} p(X)p(Y=y\left|X\right.)} \propto \max_x pX(x)pY\left|X(y,x)\right. \quad (2)$$

Where capital letters denote random quantities and lower case letters denote the deterministic realizations.

The formula seems easy. However, remember that X is 2-D image of integer values, optimization is prohibitively difficult, in general. Further, characterization of a prior p(X) is not easy, either.

Multi scale image segmentation methods and advantages: Traditionally, image segmentation has been accomplished using Markov Random Field (MRF) models. The global statistical models in the MRF theory substantially improve over local methods[15,16]. They provides a powerful framework for studying

nonlinear interactions between different features[16]. Under MAP criterion, they lead to the minimization of a global energy function which is very computationally expensive[15,17].

The Multi-Scale Hidden Markov model (MHMM) proposed in[18] is used to deal with this problem in this study. The states (an edge or smooth area), of the wavelet coefficients at different scales are modeled as Markov chain.

According to results and conclusions from other literature[8,20], this kind of multi-scale modeling not only captures the key inter-scale physical dependency present in natural signals and images (of course, the effect is demonstrated by real applications), it also leads to computationally efficient (usually scale-recursive) algorithms.

Some general peoperties of screte wavelet transform: In general, multi-scale analysis refers to the study of behavior of signals or images at various spatial and/or temporal resolutions[21]. The wavelet transform has several attractive properties which allow tremendous applications for signal and image processing[18]. The first two are multiresolution and locality: Each wavelet $\psi_k^j(t)$ is only a dilated and translated version of original mother wavelet $\psi(t)$ and is localized simultaneously in time and frequency. The 3rd one is compressive property: The wavelet coefficients of real-world signals/images tend to be sparse. The 4th and 5th properties are clustering and persistence: If one wavelet coefficient is large/small, then its adjacent coefficients are very likely to also be large/small and large/small values of wavelet coefficients tend to propagate across scales. The clustering property suggests that coefficients may have strong dependencies within scale, while persistence leads to dependencies of wavelet coefficients across scale. The hidden Markov model used in section (3.4) is to try to utilize these properties.

Theory and method -image segmentation by multi-scale hidden markov model of wavelet coefficients: The main ideas originate from[18]. This method consists of two procedures: The first one is edge detection and the other is label (state) estimation. The idea of applying wavelet analysis to edge detection is quite simple. Roughly speaking, wavelet coefficients represent the differences between function approximations at different scales (or resolutions), one kind of differentiation, intuitively, is well suited for edge detection. In the following four subsections, I' ll address in detail how to achieve the first step: edge-detection. Then I'll explain briefly the second step and how to apply our formulations in 1-D case to 2-D image.

Likelihood function for the fMRI data: Modelling the spatial correlation of the fMRI data in making statistical inference is a challenging problem for which not many solutions have been proposed than the solutions addressing the temporal dependencies. In this study, we deal with fMRI magnitude images. As noted in the beginning, the fMRI correlation image in this case may be modeled as 2-d Gaussian process. The correlation between reference, which is characteristic of the BOLD response[1] and assumed known in this study and the magnitude time series $z: c = r^T z = \sum_{j=1}^{N} r_j z_j$ is Gaussian distributed (provide the signal noise ratio is not very small). But for simplicity, let us first consider the one-dimensional analog. Yielding to invention, we assume that the length of the correlation sequence is a power of 2. The observation model is:

$$c_k^J = p_k^J + w_k^J, k = 0,...,2^J - 1 \quad (3)$$

Where:

$c^J = \{c_k^J\}$ = Observations (spatial correlations)

$\rho^J = \{\rho_k^J\}$ = True correlation values

$\{w_k^J\}$ = Noise

Now we are going to use a special (the simplest) multi-scale analysis, i.e., Haar wavelet transform on the data:

$$c_k^j = \frac{c_{2k}^{j+1} + c_{2k+1}^{j+1}}{\sqrt{2}}, k = 0,...,2^j - 1, j_0 \le j \le J - 1$$

The multi-scale analysis of the data ρ and w is defined in an analogous way.

It is then straightforward to see that:

$$c_k^j = \rho_k^j + w_k^j$$

The noise $\{w_k^J, k = 0,...,2^J - 1\}$ are assumed to be (spatially) independent, identically distributed zero-mean Gaussian random variables with variance σ^2. It then follows that the preceding sentence is also true for any j, resulting in the likelihood function:

$$p\left(c^j \middle| \rho^j\right) = \prod_{k=0}^{2^{j-1}} N\left(c_k^j \middle| \rho_k^j, \sigma^2\right), \quad j_0 \le j \le j \quad (4)$$

Where, $c^j \equiv \left\{c_k^j\right\}_{k=0}^{2^{j-1}}$ and similarly for $\rho^j, N(x|\rho,\sigma^2)$ denotes a Gaussian density with mean ρ and variance 2 evaluated at the point x.

The relationship between a "parent" (e.g., c_k^j) and a "child" (e.g., c_{2k}^{j+1}) is very important in multi-scale data

analysis. The parent-child conditional likelihood in our case turns out to be:

$$p(c_{2k}^{j+1} \mid c_{k,\rho}^{j}) = N\left(c_{2k}^{j+1} \mid \frac{c_k^j}{\sqrt{2}} + \frac{\theta_k^j}{\sqrt{2}}, \frac{\sigma^2}{2}\right) \quad (5)$$

Where, the canonical parameter:

$$\theta_k^j = \frac{\rho_{2k}^{j+1} - \rho_{2k+1}^{j+1}}{\sqrt{2}} \quad (6)$$

Is simply the Haar wavelet coefficient of true correlation ρ at scale j and location k. This nice form of the likelihood suggests the use of so called conjugate prior for the wavelet coefficients in the following subsection, which complements the observation model and leads to closed form for the posterior of the states.

Further, the likelihood function (4) with j = J can be factorized as follows[2]:

$$p(c\left|\rho\right.) = p(c^{J0}\left|\rho^{J0}\right.)\prod_{j=J_0}^{J-1}\prod_{k=0}^{2^{j-1}} p(c_{2k}^{j+1} \mid c_k^j, \theta_k^j) \quad (7)$$

Where, J_0 is the coarsest scale for the analysis (usually we use $J_0 = 0$), $p(c_{2k}^{j+1}\left|c_k^j, \theta_k^j\right.)$ is given by (5) and $p(c^{J0}\left|\rho^{J0}\right.)$ is given by (4) with $j = J_0$.

Multi-scale hidden Markov model (MHMM) for the prior of the wavelet coefficients: Now let's consider prior (joint) probability for the (unknown) wavelet coefficients θ. A simple approach is to model them as independent Gaussian mixture random variables. We move beyond this simple prior, by specifying probabilistic dependencies between the states underlying the mixtures of parent and child wavelet coefficients. To deduce discrete state estimations from continuous data, the key point for our algorithm is to associate the continuous wavelet coefficients with a 2-state discrete Markov chain.

Specifically, for our real problem, the states (edge or smoothness) of the Markov chain are unknown ("hidden") and represent the presence or absence of edges: state 0 indicates a homogenous region, state 1 represents the existence of an edge. We perceive that the under lying signal is generally smooth with a few large edges, then the following modeling is intuitively reasonable, i.e., we consider two states mixture model where states '0' is a highly probable low-variance Gaussian density, indicative of a homogenous region, while states '1', corresponding to another less likely Gaussian density with a large variance, indicates the presence of an edge (no-smooth area). Using this interpretation, we may test for the presence of an edge simply by checking whether following condition holds or not:

$$p(s_k^j = 1\left|c\right.) > p(s_k^j = 0\left|c\right.) \quad (8)$$

If it holds, we conclude there is an edge at scale j and location k. Otherwise, there is not.

Mathematically, the MHMM is based on the assumption that the value of each states s_k^j is caused by the value of its parent state. This leads to the factorization of the joint state probability function:

$$p(s) = \prod_{j=J_0}^{J-1}\prod_{k=0}^{2^{j-1}} p(s_k^j \mid s_{\lfloor k/2 \rfloor}^{j-1}) \quad (9)$$

Where, $p(s_0^0\left|s_0^{-1}\right.) \equiv p(s_0^0)$. At the coarsest scale j = 0, we have no parent wavelet coefficients and so we introduce a prior for the states of the wavelet coefficients $p(s_0^0)$. Note that ρ_0^0 is the global average correlation data.

Another property of the HMM is that, given their respective states values, all parameters θ are conditionally independent[19]. That is:

$$p(\theta \mid s) = \prod_{J=J_0}^{J-1}\prod_{k=0}^{2^{j-1}} p(\theta_k^j \mid s_k^j) \quad (10)$$

Where, the prior probability $p(\theta_k^j\left|s_k^j\right.)$ is assumed to be Gaussian

$$p(\theta_k^j\left|s_k^j\right. = m) = N(\theta_k^j\left|\mu_m^j, \tau_m^{j2}\right.) \quad (11)$$

We regard the signal and its wavelet coefficients as realizations from a large family of random signal. Therefore, collectively, we assume $\mu_m^j = 0$

Solution for joint a posterior state probability: Having setup the formulations for likelihood and a prior, we are now ready to determine the a. posterior density of the joint states given observation. Note that:

$$\begin{aligned} p(s \mid c) &= \int p(s, \theta \mid c) d\theta \\ &\propto \int p(c \mid s, \theta) p(\theta \mid s) p(s) d\theta \\ &= \prod_{j=J_0}^{J-1}\prod_{k=0}^{2^{j-1}} \int p(c_{2k}^{j+1} \mid c_k^j, \theta_k^j, s_k^j) p(\theta_k^j \mid s_k^j) p(s_k^j \mid s_{\lfloor k/2 \rfloor}^{j-1}) d\theta_k^j \\ &= \prod_{j=J_0}^{J-1}\prod_{k=0}^{2^{j-1}} p(s_k^j \mid s_{\lfloor k/2 \rfloor}^{j-1}) L_k^j(s_k^j) \end{aligned} \quad (12)$$

Where m_k^j is one particular state value assumed by random variable s_k^j and $L_k^j(s_k^j = m) \propto p(c_{2k}^{j+1} | c_k^j, s_k^j = m)$, the essential ingredients for our estimation of the a posterior states, are actually marginal likelihoods. From the likelihood function in (5) and the prior in (1), we drive them to be:

$$L_k^j(m) = \int p(c_{2k}^{j+1} | c_k^j, \theta_k^j) p(\theta_k^j | s_k^j = m) d\theta_k^j = N\left(c_{2k}^{j+1} | \frac{\mu^j m}{\sqrt{2}} + \frac{c_k^j}{\sqrt{2}}, \frac{\tau_m^{j2} + \sigma^2}{2}\right) \quad (13)$$

$$J_0 \le j \le J-1, k = 0,1,\ldots,2^j - 1, m = 0,1$$

Proof:

$$y_{2k}^{j+1} = \frac{y_k^j + \theta y_k^j}{\sqrt{2}} = \frac{y_k^j}{\sqrt{2}} = \frac{\theta_k^j}{\sqrt{2}} + \frac{\theta w_k^j}{\sqrt{2}}$$

where the first two terms are independent of the third. Further, recall:

$$p(\theta_k^j | s_k^j = m) = N(\theta_k^j | \mu_m^j, \tau_m^{j2}), \text{ i.e., } \theta w_k^j \sim N(0, \sigma^2), \theta_k^j | s_k^j = m \sim N(\mu_m^j, \tau_m^{j2})$$

Therefore,

$$E(y_{2k}^{j+1} | y_k^j, s_k^j = m) = \frac{\mu_m^j}{\sqrt{2}} + \frac{y_k^j}{\sqrt{2}}$$

$$Var(y_{2k}^{j+1} | y_k^j, s_k^j = m) = \frac{\tau_m^{j2} + \sigma^2}{2}$$

So, $L_k^j(m) p(y_{2k}^{j+1} | y_k^j, s_k^j = m)$ (where y_k^j is regarded as constant), one marginal density function, has the closed form representation in Eq. 12.

Marginal a posteriori state probability calculation: With these formulations ready, we can use the upward-downward algorithm[18] to determine the most likely marginal a posteriori state for the wavelet coefficients θ_k^j and then use Eq. 8 to test the presence of an edge.

In the upward-downward algorithm, the Up step marginalizes the joint posteriori state probability recursively from the finest scale j = J-1 to the coarsest scale j = 0. At the end the posterior state probabilities $\{p(s_0^0 = m | c)\}_{m=0}^{M-1}$ are provided and partial marginalizations are also stored for use in the Down Step. The Down Step computes the marginal posterior state probabilities for each s_k^j recursively. For the specific flow of the upward-downward algorithm, refer to[18].

Segmentation: After the edges are determined, it straightforward to formulate likelihood ratio test to estimate the label (state) of each homogenous region.

Consider the following multi-hypothesis problem. The observation $c = [c_1, c_2 \ldots c_n]^T$ within each homogenous region is Gaussian random vector of dimension n. The M hypothesis are:

$$H_i: c \sim N(m_i, c_i), i = 1,2,\ldots,M \quad (14)$$

Where m_i and C_i is the mean vector and covariance matrix of the observation under the ith i = 1,2,…,M hypothesis, which are assumed to be known. Suppose each hypothesis equally likely and minimum error criterion is adopted, the decision rule then boils down to hosing H_j where:

$$j = \arg\min_i \left\| c - m_i \right\|^2 + \ln \left| c_i \right| \quad (15)$$

where $\left\| c - m_i \right\|^2 \equiv (c - m_i)^T C^{-1} (c - m_i)$ and $\left| c_i \right|$ and C_i.

In the following simulated processing, the observation within each homogenous region is assumed to be independent and identically distributed (i. i. d.), that is, I assume $m_i = m_i 1$ and $C_i = \sigma_i^2$. Some important practical problems are: How to assign the a prior probability for s_0^0, how to assign transition probabilities for the states:

$$p(s_k^j = m | s_{\lfloor k/2 \rfloor}^{j-1} = m') \quad (16)$$

and how to determine the parameters for the Gaussian distribution characterizing each homogeneous region. Theoretically, these parameters are estimated by a complicated E-M algorithm. For our initial investigation, we set them empirically (by observation). It turns out that our experiment results are insensitive to the a priori probability and transition probability. The robustness is a nice feature.

Extension to 2 dimensions: We can extend the multi-scale analysis and MHMMs easily from 1-D sequence to 2-D images. Instead of taking the usual 2-D wavelet transform to the original image; we use the following conversion method. First we convert the original 2-D image into 1-D sequence and then apply previous 1-D wavelet analysis to the resulting sequence. The conversion details are: First split the image vertically into two halves, then horizontally splitting each half into two quarters, and iterate until each one is a 1×1 pixel. The merit of this conversion is that it retains the original spatial configuration. Refer to Fig. 1 for details.

One simulation testifying image segmentation by our algorithm: Figure 2a and b shows a simulated noisy image, the grey level of the segmented image respectively. A two-state MHMM was specified for this problem with the following parameter settings:

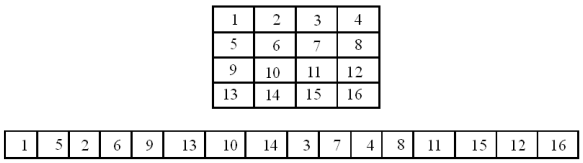

Fig. 1: Conversion of a 2-D image to 1-D sequence

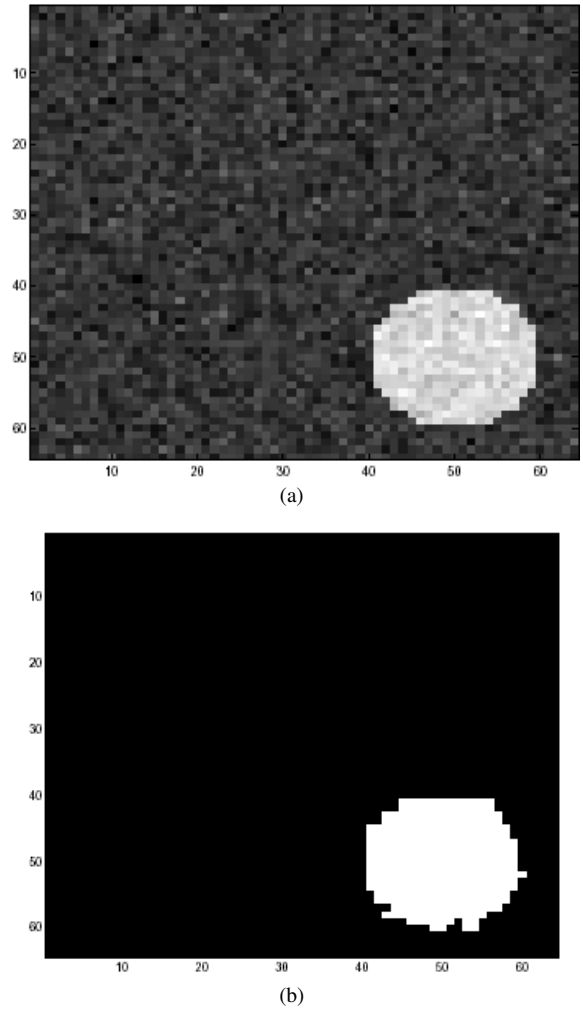

Fig. 2: (a) Noisy image, (b) Segmented image

$$\tau_0 = 1,$$
$$\tau_1 = 100,$$
$$p_0^0(0) = .9,$$
$$p_k^j(0|0) = 0.9, k = 0,...,2^j - 1, j = 1,...,J - 1,$$
$$p_k^j(0|1) = 0.25, k = 0,...,2^j - 1, j = 1,...,J - 1.$$

Figure 2b demonstrates that the results are excellent.

RESULTS

Processing results for fmri simulated data by two-step approach as stated in the beginning, the method for fMRI detection in this study involves a two-step procedure: Multiscale image segmentation will be first used to break the correlation image into different regions of homogeneous statistical behavior, each region will then be tested independently as active or inactive by single pixel detection method.

In order to see the potential of this method for fMRI detection, the following experiment is conducted to compare results from the combined effects of single pixel detection and image segmentation with results obtained based solely on pixel-wise detection.

Using the model in Eq. 1 of[1], a simulated fMRI complex time series is generated at each pixel. In order to simulate the profile of the brain, the magnitudes of the baseline signal (a's in Eq. (1) in[1] in the complex time series roughly follow the magnitude data from a static brain image. Actually the original complex data used in this example are exactly the same as Fig. 3b in[1], which is reproduced here as Fig. 3a for sake of comparison, Next, the correlation value at each pixel is computed by correlating the magnitude time series with the reference to produce Fig. 3b.

Figure 3c is the segmented result of correlation in Fig. 3b based on our algorithm. There are M = 2 lablels: each pixel is assigned to either 0 or 1 to its lablel. The parameters in the example are set to be:

$$\sigma^2 = 1;$$
$$\tau_0^2 = 1;$$
$$\tau_1^2 = 100;$$
$$p_0^0(0) = .95;$$
$$p(0|0) = 0.95, k = 0,...,2^j - 1, j = 1,...,J - 1;$$
$$p(0|1) = 0.05, k = 0,...,2^j - 1, j = 1,...,J - 1;$$
$$m_0 = 0;$$
$$m_1 = 2.$$

In this example, I set σ_1 and σ_0 (variances for the Gaussian characterizing two homogeneous regions) to be equal. The test criterion in Eq. 15 reduces to simple from in this case. The original simulated active region in Fig. 3a is a 9*9 square; in Fig. 3c the white region is 10*8 rectangle. They are in good agreement but not in prefect match. This is not surprising, since, in general, we cannot guarantee the segmentation step produces exactly the same geometry as the original simulated regions: What is shown here is just one realization of many random simulations.

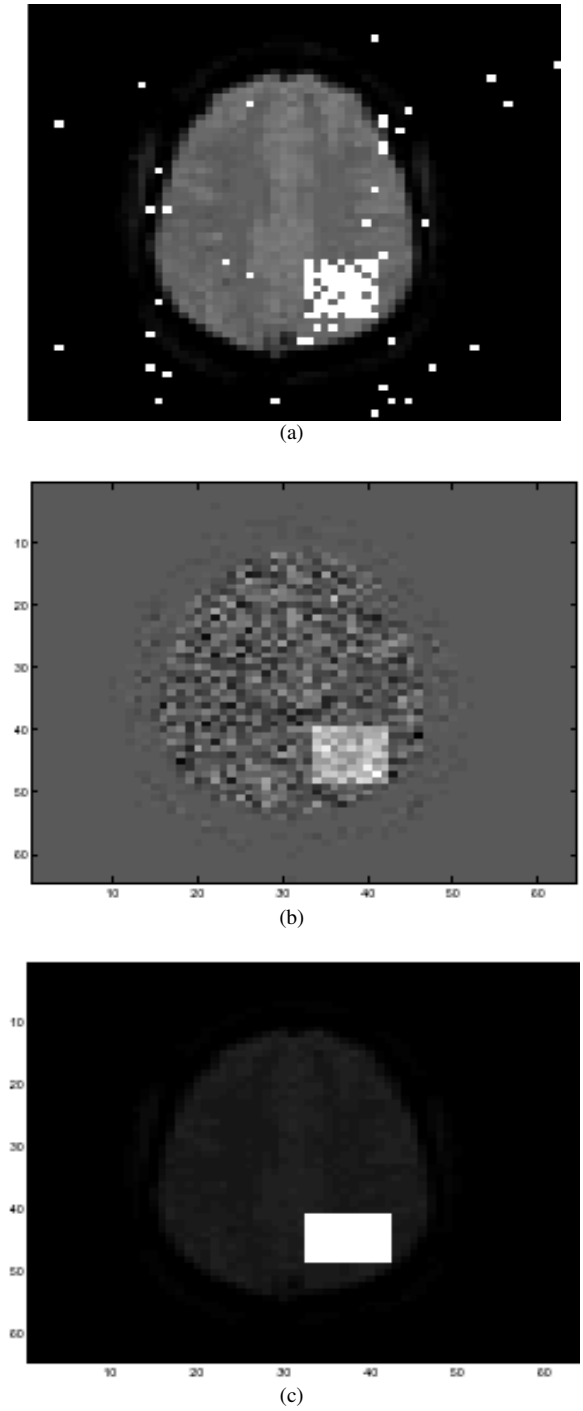

Fig. 3: Comparison of detection results from one step pixelwise method and from two-step approach. (a) Detection image from[1] (b) fMRI correlation image, (c) Segmented image of (b), also final detection results by combination use of image segmentation and single pixel detection

Next I consider applying single pixel detection technique in[1] to each of the above homogeneously (statistically) distributed region. The idea is to regard each homogeneous region as one large, macro-pixel: The average of all time series inside each macro-pixle is taken to be the new time series characterizing this macro-pixel, then apply single pixel detection method (magnitude correlation method in[1]) to the new time individually to determine which of these macro-pixels is active and which one is in active. By this approach, the macro-pixels (original pixels in Fig. 3b), in contrast to macro-pixels) corresponding to the black region in Fig. 3c are all inactive, which is expected since this region contains a large outside the brain. The micro-pixels in Fig. 3b corresponding to the white region in Fig. 3c turns out to be all active. In other words, by this approach, only 9 pixels inside the square are missed while the 8 pixels outside the square false-alarmed.

Now let us take a comparison between Fig. 3a and 3c in this study. For the former, we see spurious activation regions outside the brain. However, the falsely alarmed regions disappear in Fig. 3c (except for 8 pixels outside the square) after combining image segmentation with single pixel detection. Pure single pixel detection methods failed to detect some active pixels inside the small square in last figure of[1]. However, these regions (except for 9 pixels) are now correctly detected by combining image segmentation with single pixel detection.

The enhancement of detection efficiency is clearly visible and also easily understandable. Actually we are given spatial-temporal series. However, the pixel-wise detection method only takes temporal information into account: Spatial information is completely ignored. The image segmentation algorithm in this study exactly complements the pixel-wise detection and remedies its shortcoming: it utilizes the spatial correlation information inherent in the data. So it is no wonder that the detection performance improves after image segmentation.

One point to be noted is that in Fig. 3c the brain profiles are overlapped, as are the case in the last figure in[1].

Our simulations were done using Matlab 7.0 software and on a usual PC with Windows XP (main frequency 2GHz, memory 1GB, hard disk 80 GB).

DISCUSSION

This paper put forward a novel two-step procedure (image segmentation combined with single pixel-wise detection) for fMRI detection. The image segmentation (or edge detection) using a novel hidden Markov

modeling wavelet coefficients actually exploits the spatial information of fMRI correlation image, while single pixel-wise detection utilizes the temporal information in fMRI time series.

Considering utilizing spatial correlation information, this study uses Bayesian image segmentation method. Other approaches involving spatial consideration can be used as well. For example, clustering analysis are gaining more recognition in this field[23]. Formulation as decentralized detection problem is also a possible candidate[27,24].

CONCLUSION

FMRI signals are actually both temporally and spatialy dependent. Pixel-wise detection, however, considers only temporal correlation information and ignores spatial correlation information. In order to remedy this deficiency, this study uses a multi-scale image segmentation algorithm to first segment an fMRI correlation image into several regions, each with homogeneous statistical behavior. A single pixel detection algorithm is the applied to each homogeneous region. Extensive simulations demonstrate improved efficacy of our method.

REFERENCES

1. Nan, F. and R.D. Nowak, 1999. Generalized likelihood ratio detection for fMRI using complex data. IEEE Trans. Med. Imag., 18: 320-329. DOI: 10.1109/42.768841.
2. Bandettini, P.A., A. Jesmanowvicz, E.C. Wong and J.S. Hyde, 1993. Processing strategies for time coarse data sets in functional MRI of the human brain. Magn. Reson. Med., 30: 161-173. DOI: 10.1002/mrm.1910300204.
3. Genoves, C.R., D. Cnol and W. Fedy, 1977. Estimating test-retst reliability of functional MR imaginig: Statistical methodlogy. Magn. Reson. Med., 38: 497-507.
4. Siewrt, B., B.M. Bly, G.S. Hlaug, D.G. Darby, V. Thangarj, S. Warach and R. Edelman, 1996. Comparison of the bold and epistar-technique for functional brain imaging using signal detection theory. Magn. Res. Med., 36: 249-25. DOI: 10.1002/mrm.1910360212.
5. Golay, X., S. Kolias, G. Stol, D. Meir, A. Valvanis and P. Boesiger, 1998. A new correlation-based fuzzy logic clustering algorithm for fMRI, Magn. Res.Med., 40: 249-260. DOI: 10.1002/mrm.1910400211.
6. Keown, M.J.M., S. Makeig, G.G. Brown, T.P. Jung, S. Kinderman, A.J. Bel and T.J. Sejnowski, 1998. Analysis of MRI data by blind separation into independent spatial components. Hum. Brain Map., 6: 160-188. DOI: 10.1002/(SICI) 1097-0193 (1998)6:3<160:: AID-HBM5>3.0.CO;2-1.
7. Frank, L.R., R. Buxton and E.C. Wong, 1998. Probabilistic analysis and functional magnetic resonance imaging data. Magn. Reson. Med., 39: 132-148. http:// cat.inist.fr/?aModele=afficheN&cpsidt=2090834.
8. Bouman, C. and M. Shapiro, 1994. A multi scale random field model for Bayesian image segmentation. IEEE Trans. Image Proc., 3: 162-177. DOI: 10.1109/83.277898.
9. Fadil, M.J. and E.T. Bulmore, 2004. A comparative evaluation of wavelet-based methods for hypothesis testing of brain activation maps. NeuroImage, 23: 1112-1128. DOI: 10.1016/j.neuroimage.2004.07.034.
10. Long, C., E.N. Brown, D. Manoach and V. Solo, 2004. Spatiotemporal wavelet analysis for functional MRI. NeuroImage, 23: 500-516. DOI: 10.1016/j.neuroimage.2004.04.017.
11. Bulmore, E., J. Fadil, V. Maxim, L. Sendur, B. Whitcher, J. Suckling, M. Brammer and M. Breakspear, 2004. Wavelets and functional magnetic resonance imaging of the human Brain. NeuroImage, 23: 234-249. DOI: 10.1016/j.neuroimage.2004.07.012
12. Desco, M., J.A. Hernadez and M. Brammer, 2001. Multi resolution analysis in fMRI: Sensitivity and specificity in the detection of Brain activation. Human Brain Map., 14: 16-27. DOI: 10.1002/hbm.1038.
13. Friston, J., A.P. Holmes, K.J. Worsley, J.B. Poline, C. Frith and R.S.J. Frackowiak, 1995. Statistical parametric maps in functional imaging: A general linear approach. Human Brain Map., 2: 189-210. DOI: 10.1002/hbm.460020402.
14. Cox, Robert W., AFNI: Software for analysis and visualization of functional magnetic resonace neuro images. 1996. Comput. Biomed. Res., 29: 162-73. DOI: 10.1002/hbm.460020402.
15. Geman, S. and D. Geman, 1984. Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. IEEE Trans. Patt. Anal. Mach. Intel., 6: 721-741. DOI: 10.1080/02664769300000058.

16. Perez, P. and F. Heitz, 1996. Restriction of Markov random field on a graph and multiresolution Statistical image modeling. IEEE Trans. Inform. Theory, 42: 180-190. DOI: 10.1109/18.481788.
17. Malfait, M. and D. Rose, 1997. Wavelet-based image denoisng using a Markov random field a priori model. IEEE Trans. Image Process., 6: 549-565. DOI: 10.1109/83.563320.
18. Crouse, M.S., R.D. Nowak and R.G. Barniuk, 1998. Wavelet based statistical signal processing using hidden Markov models. IEEE Trans. Signal Process., 46: 886-902. DOI: 10.1109/78.668544.
19. Rabiner, L., 1989. A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE, 77: 257-285. DOI: 10.1109/5.18626.
20. Luettgen, M.R., W.C. Karl, A.S. Willsky and R.R. Tenney, 1993. Multiscale representations of Markov Random fields. IEEE Trans. Signal Process., 41: 3377-3396. DOI: 10.1109/78.258081.
21. Rioul, O. and M. Vetrli, Wavelets and signal processing. IEEE Signal Process. Magazine, 10: 14-38. DOI: 10.1109/79.91217.
22. Frank, L.R., R.B. Buxton and E.C. Wong, 1998. Probabilistic analysis and functional magnetic resonance imaging data. Magn. Reson. Med., 39: 132-148. DOI: 10.1002/mrm.1910390120.
23. Golay, X., S. Kolias, G. Stol, D. Meir, A. Valvanis and P. Boesiger, 1998. A new correlation based fuzzy logic clustering algorithm for fMRI. Magn. Resonance Medicine, 40: 249-260. DOI: 10.1002/mrm.1910400211.
24. Polychronpouls, G. and J.N. Tsitklis, 1990. Explicit solutions for some simple decentralized problems. IEEE Trans. Aerospace Elect. Syst., 26: 282-292. DOI: 10.1109/7.53441.
25. Desai, M., R. Mangoubi, J. Shah, W. Karl, H. Pien, A. Worth and D. Kenedy, 2002. Functional MRI activity characterization using response time shift estimates from curve evolution. IEEE Trans. Med. Imag., 21: 1402-1412. DOI: 10.1109/TMI.2002.806419.
26. Desai, M. and R. Mangoubi, 2003. Robust gaussian and non-gaussian matched subspace detection. IEEE Trans. Signal Process., 51: 3115-3127. DOI: 10.1109/TSP.2003.818907.
27. Irving, W.W. and J.N. Tsitklis, 1994. Some properties of optimal thresholds in decentralized detection. IEEE Trans. Automatic Control, 39: 835-838. DOI: 10.1109/9.286264.
28. Zadeh, G.A.H., B.A. Ardekani and H.S. Zadeh, 2003. Activation detection if MRI using a maximum energy ratio statistic obtained by adaptive spatial filtering. IEEE Trans. Med. Imag., 2: 795-805. DOI: 10.1109/TMI.2003.815074.
29. Rowe, D.B. and B.R. Logan, 2004. A complex way to compute fMRI activation. NeuroImage, 23: 1078-1092. DOI: 10.1016/j.neuroimage.2004.06.042.
30. Rowe, D.B. and B.R. Logan, 2005. Complex fMRI analysis with unrestricted phase is equivalent to a magnitude only model. NeuroImage, 24: 603-606. DOI: 10.1016/j.neuroimage.2004.09.038.
31. Mallat, S. and S. Zhong, 1992. Characterization of signals from multiscale edges. IEEE Trans. Patt. Anal. Mach. Intel., 14: 710-732. DOI: 10.1109/34.142909.