NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 27 (2005) 368 – 376

A simple nonparametric statistical thresholding for MEG spatial-filter source reconstruction images

Kensuke Sekihara,[a,*] Maneesh Sahani,[b] and Srikantan S. Nagarajan[c]

[a]Department of Systems Design and Engineering, Tokyo Metropolitan University, Asahigaoka 6-6, Hino, Tokyo 191-0065, Japan
[b]Gatsby Computational Neuroscience Unit, University College London, 17 Queen Square, London, WClN 3AR, UK
[c]Department of Radiology, University of California, San Francisco, 513 Parnassus Avenue, S362 San Francisco, CA 94143, USA

Received 12 October 2004; revised 3 February 2005; accepted 5 April 2005
Available online 29 June 2005

This paper proposes a simple statistical method for extracting target source activities from spatio-temporal source activities reconstructed from MEG measurements. The method requires measurements in a control condition, which contains only non-target source activities. The method derives, at each pixel location, an empirical probability distribution of the non-target source activity using the time course reconstruction obtained from the control period. The statistical threshold that can extract the target source activities is derived from the empirical distributions obtained from all pixel locations. Here, the multiple comparison problem is addressed with a two-step procedure involving standardizing these empirical distributions and deriving an empirical distribution of the maximum pseudo *T* value at each pixel location. The results of applying the proposed method to auditory-evoked measurements are presented to demonstrate the method's effectiveness.

Keywords: Magnetoencephalography; Statistical threshold; Biomagnetism; Nonparametric statistics; Spatial filter; Spatio-temporal reconstruction

Introduction

Among the various technologies for noninvasive neural measurement, the major advantage of magnetoencephalography (MEG) is its ability to provide fine temporal resolution on the order of milliseconds (Hämäläinen et al., 1993). Neuromagnetic imaging can therefore visualize neural activities with such a fine time resolution and provide functional information about brain dynamics. One major problem is that the measured MEG signal generally contains not only a magnetic field associated with the signal sources of interest but also contains interference magnetic fields generated from non-target activities. Such non-target activities include spontaneous brain activities or some evoked activities that are not the interest of the current investigation. These non-target activities generally overlap with the target signal activities in the source reconstruction, and they often make interpreting the reconstruction results difficult.

In most studies using positron emission tomography (PET) or functional magnetic resonance imaging (fMRI), the experiments are carefully designed to extract only target activities. A common example of such experimental design contains two kinds of stimuli: a task stimulus and a control stimulus. The task stimulus generally elicits the target cortical activities as well as other activities associated with the target activities. The control stimulus is designed to elicit only the latter activities. Then, by calculating the statistical difference between the images measured with the two kinds of stimuli, the target activities can be revealed. Both parametric (Friston et al., 1995; Worsley et al., 1996) and nonparametric statistics (Nichols and Holmes, 2001) have been used to calculate such statistical differences.

This paper proposes a simple method of statistical subtraction between task and control measurements. The method is applicable to spatio-temporal source reconstruction from MEG/EEG measurements. It assumes neural activities to be quasi-stochastic, and it uses nonparametric statistics to derive an empirical probability distribution of these activities using the time course reconstruction in the control period. This empirical distribution is then used for deriving an appropriate value for the statistical thresholding. The thresholding can extract the target activities that exist only in the task measurements by eliminating other non-target activities that exist both in the task and control measurements.

In this paper, we present the proposed statistical thresholding method using spatial filter source reconstruction (Sekihara and Nagarajan, 2004). This is because the formulation of the spatial filter is relatively simple and the spatial filter techniques have been successfully applied to MEG source analysis (Hashimoto et al., 2003; Ishii et al., 2003). However, the applicability of the proposed method is not limited to the spatial filter formulation and it can be used with any type of source estimation method that can provide the spatio-temporal source reconstruction, i.e., that can reconstruct source time courses at all pixel locations.

* Corresponding author. Fax: +81 42 585 8642.
E-mail address: ksekiha@cc.tmit.ac.jp (K. Sekihara).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2005.04.009

Spatial filter formulation for MEG source reconstruction

Definitions

We define the magnetic field measured by the mth detector coil at time t as $\boldsymbol{b}(t)$ and a column vector $\boldsymbol{b}(t) = [b_1(t), b_2(t), \ldots, b_M(t)]^T$ as a set of measured data where M is the total number of sensor coils and superscript T indicates the matrix transpose. The spatia1 location is represented by a three-dimensional vector $\boldsymbol{r}$: $\boldsymbol{r} = (x, y, z)$. The second-order moment matrix of the measurement is denoted $\mathbf{R}$, i.e., $\mathbf{R} = \langle \boldsymbol{b}(t)\boldsymbol{b}^T(t)\rangle$ where $\langle\bullet\rangle$ indicates the ensemble average, which is replaced by the time average over a certain time window in practice. (When $\langle \boldsymbol{b}(t)\rangle \approx 0$ holds, $\mathbf{R}$ is also equal to the covariance matrix of the measurement.) The magnitude of the source moment is denoted $s(\boldsymbol{r}, t)$. The source orientation is defined as a three-dimensional column vector $\boldsymbol{\eta}(\boldsymbol{r}) = [\eta_x(\boldsymbol{r}), \eta_y(\boldsymbol{r})\ \eta_z(\boldsymbol{r})]^T$ whose ζ component (where ζ equals x, y, or z) is equal to the cosine of the angle between the direction of the source and the ζ axis.

We define $\boldsymbol{l}_m^\zeta(\boldsymbol{r})$ as the output of the mth sensor; the output is induced by the unit-magnitude source located at $\boldsymbol{r}$ and pointing in the ζ direction. The column vector $\boldsymbol{l}_\zeta(\boldsymbol{r})$ is defined as $\boldsymbol{l}_\zeta(\boldsymbol{r}) = [l_1^\zeta(\boldsymbol{r}), l_2^\zeta(\boldsymbol{r}), \ldots, l_M^\zeta(\boldsymbol{r})]^T$. We define the lead field matrix, which represents the sensitivity of the whole sensor array at $\boldsymbol{r}$, as $\mathbf{L}(\boldsymbol{r}) = [\boldsymbol{l}_x(\boldsymbol{r}), \boldsymbol{l}_y(\boldsymbol{r}), \boldsymbol{l}_z(\boldsymbol{r})]$.

Typical non-adaptive and adaptive spatial filters

Spatial filter techniques estimate the source current density by applying a simple linear operation to the measured data, i.e.,

$$\hat{s}(\boldsymbol{r}, t) = \boldsymbol{w}^T(\boldsymbol{r})\boldsymbol{b}(t) = \sum_{m=1}^{M} w_m(\boldsymbol{r})b_m(t), \tag{1}$$

where $\hat{s}(\boldsymbol{r}, t)$ is the estimated source magnitude. The column vector

$$\boldsymbol{w}(\boldsymbol{r}) = [w_1(\boldsymbol{r}), \ldots\ldots, w_M(\boldsymbol{r})]^T$$

represents a set of the filter weights. The filter weights $\boldsymbol{w}(\boldsymbol{r})$ should only pass the signal from a source with a location $\boldsymbol{r}$ and reject the signals generated at other locations. Since the source is a three-dimensional vector quantity, the weight vector generally depends on the source directions. In this paper, we define $\boldsymbol{w}(\boldsymbol{r})$ as the weight vector in the optimum direction $\boldsymbol{\eta}_{\text{opt}}(\boldsymbol{r})$, which is determined as the direction that gives the maximum spatial-filter outputs at each $\boldsymbol{r}$.

There are two types of spatial filter techniques. One is a non-adaptive method in which the filter weight is independent of the measurements. The best-known non-adaptive spatial filter is the minimum-norm estimate (Hämäläinen and Ilmoniemi, 1984). The filter weight is expressed as

$$\boldsymbol{w}(\boldsymbol{r}) = \boldsymbol{G}^{-1}\boldsymbol{L}(\boldsymbol{r})\boldsymbol{\eta}_{\text{opt}}(\boldsymbol{r}). \tag{2}$$

The matrix $\mathbf{G}$ is often referred to as the gram matrix, which is given by calculating the overlap between the lead fields,

$$\mathbf{G} = \int \boldsymbol{L}(\boldsymbol{r})\boldsymbol{L}^T(\boldsymbol{r})d\boldsymbol{r}. \tag{3}$$

The estimated current density is then expressed as

$$\hat{s}(\boldsymbol{r}, t) = \boldsymbol{\eta}_{\text{opt}}^T(\boldsymbol{r})\boldsymbol{L}^T(\boldsymbol{r})\mathbf{G}^{-1}\boldsymbol{b}(t). \tag{4}$$

The minimum-norm spatial filter in its original form is known to give erroneous source reconstruction results. Therefore, to improve the performance, it is often used with some kind of constraint on source distributions (Baillet et al., 2001; Dale et al., 2000; Pascual-Marqui, 2002).

The other type of spatial filter is an adaptive spatial filter in which the filter weight depends on the measurements. The best-known adaptive spatial filter is the minimum variance spatial filter, which is customarily referred to as the minimum-variance beamformer (Robinson and Vrba, 1999; Sekihara and Seholz, 1996; Sekihara et al., 2001; van Veen et al., 1997). In this method, the weight vector is given by

$$\boldsymbol{w}^T(\boldsymbol{r}) = \frac{\boldsymbol{l}^T(\boldsymbol{r})\boldsymbol{R}^{-1}}{\boldsymbol{l}^T(\boldsymbol{r})\boldsymbol{R}^{-1}\boldsymbol{l}(\boldsymbol{r})}, \tag{5}$$

where $\boldsymbol{l}(\boldsymbol{r})$ is defined as $\boldsymbol{l}(\boldsymbol{r}) = \mathbf{L}(\boldsymbol{r})\boldsymbol{\eta}_{\text{opt}}(\boldsymbol{r})$. (The method for determining $\boldsymbol{\eta}_{\text{opt}}(\boldsymbol{r})$ for the adaptive spatial filter is described in (Sekihara and Seholz, 1996). Furthermore, in practice, the normalized lead field $\boldsymbol{l}(\boldsymbol{r})/\|\boldsymbol{l}(\boldsymbol{r})\|$ is used in Eq. (5) to avoid a source location bias caused by the variation of the lead field norm $\|\boldsymbol{l}(\boldsymbol{r})\|$ (Sekihara et al., 2005). The minimum-variance beamformer can be extended to the eigenspace-projection beamformer, which is known to be tolerant of errors in the forward modeling or in the estimation of the data covariance matrix (Sekihara et al., 2002). The extension is attained by projecting the weight vector in Eq. (5) onto the signal subspace of the measurement covariance matrix. That is, redefining the weight vector in Eq. (5) as $\boldsymbol{w}_{(\text{MV})}(\boldsymbol{r})$, the weight vector for the eigenspace-projection beamformer is obtained using

$$\boldsymbol{w}(\boldsymbol{r}) = \mathbf{E}_S\ \mathbf{E}_S^T\ \boldsymbol{w}_{(\text{MV})}(\boldsymbol{r}). \tag{6}$$

In this equation, $\mathbf{E}_S$ is a matrix whose columns consist of the signal-level eigenvectors of $\boldsymbol{R}$, and $\mathbf{E}_S\mathbf{E}_S^T$ is the projection matrix that projects a vector onto the signal subspace of $\boldsymbol{R}$. This eigenspace projection beamformer was used in the experiments described in Numerical experiments and Experiments.

Evaluation of statistical significance using parametric statistics

The evaluation of the statistical significance of the spatial filter outputs has typically been performed using parametric statistics (Barnes and Hillebrand, 2003; Dale et al., 2000; Gross et al., 2001; Robinson and Vrba, 1999). The basic assumption of the parametric method is that the measurement consists of deterministic signal and Gaussian noise, i.e.,

$$\boldsymbol{b}(t) = \boldsymbol{b}_I(t) + \boldsymbol{n}(t), \tag{7}$$

where $\boldsymbol{b}_I(t)$ is the signal of interest, i.e., the signal generated from brain sources that are the target of current investigation.

In Eq. (7), $\boldsymbol{n}(t)$ is the noise vector, and each element of $\boldsymbol{n}(t)$ is assumed to follow $\mathcal{N}(0,\sigma_0^2)$, which indicates the Gaussian distribution with zero mean and a variance of σ_0^2. The spatial filter outputs $\hat{s}(\boldsymbol{r}, t)$ are expressed as

$$\hat{s}(\boldsymbol{r}, t) = \boldsymbol{w}^T(\boldsymbol{r})\boldsymbol{b}(t) = \boldsymbol{w}^T(\boldsymbol{r})\boldsymbol{b}_I(t) + \boldsymbol{w}^T(\boldsymbol{r})\boldsymbol{n}(t). \tag{8}$$

Therefore, since these Gaussian processes are assumed to be uncorrelated between different sensor recordings, the outputs $\hat{s}(\boldsymbol{r}, t)$ follows $\mathcal{N}\left(\boldsymbol{w}^T(\boldsymbol{r})\ \boldsymbol{b}_I(t), \sigma_0^2\|\boldsymbol{w}(\boldsymbol{r})\|^2\right)$, which is a Gaussian distribution with a mean of $\boldsymbol{w}^T(\boldsymbol{r})\boldsymbol{b}_I(t)$ and a variance of $\sigma_0^2\|\boldsymbol{w}(\boldsymbol{r})\|^2$. Actually, since σ_0^2 must be estimated from the measured data, the distribution of $\hat{s}(\boldsymbol{r}, t)$ is not exactly represented by the Gaussian distribution but by the t distribution.

The statistical evaluation can be performed by testing the null hypothesis at each pixel location. (The null hypothesis is that there is no signal source activity.) That is, the z score under the null hypothesis, $\hat{s}(\boldsymbol{r}, t)/(\sigma_0\|\boldsymbol{w}(\boldsymbol{r})\|)$, is calculated and compared to $z_\alpha/2$, which is the two-tailed z score corresponding to the α level of significance, which is equal to the probability of Type I error. This procedure is performed at each pixel location, and if the calculated z score is higher than $z_\alpha/2$, the estimated source activity $\hat{s}(\boldsymbol{r}, t)$ is considered to be statistically significant. This procedure can be extended to incorporate the multiple comparison problem (Barnes and Hillebrand, 2003).

Proposed nonparametric statistical significance evaluation

The signal and noise model expressed in Eq. (7) is, in general, insufficient to express real-world measurements, and the measured data should be expressed as

$$\boldsymbol{b}(t) = \boldsymbol{b}_I(t) + \boldsymbol{b}_\xi(t) + \boldsymbol{n}(t), \tag{9}$$

where $\boldsymbol{b}_\xi(t)$ is the magnetic field generated from sources other than the signal sources, such as spontaneous brain activities or some evoked activities that are not the target of the current investigation. This $\boldsymbol{b}_\xi(t)$ is often referred to as the brain noise. Here, we propose a simple nonparametric method that can take such brain noise into consideration.

The spatial filter outputs obtained from $\boldsymbol{b}(t)$ is expressed as

$$\begin{aligned}\hat{s}(\boldsymbol{r},t) &= \boldsymbol{w}^T(\boldsymbol{r})\boldsymbol{b}_I(t) + \boldsymbol{w}^T(\boldsymbol{r})\boldsymbol{b}_\xi(t) + \boldsymbol{w}^T(r)\boldsymbol{n}(t)\\ &= \hat{s}_I(\boldsymbol{r},t) + \hat{s}_c(\boldsymbol{r},t),\end{aligned} \tag{10}$$

where

$$\hat{s}_I(\boldsymbol{r},t) = \boldsymbol{w}^T(\boldsymbol{r})\boldsymbol{b}_I(t) \text{ and } \hat{s}_c(\boldsymbol{r},t) = \boldsymbol{w}^T(\boldsymbol{r})(\boldsymbol{b}_\xi(t) + \boldsymbol{n}(t)). \tag{11}$$

Here, $\hat{s}_I(\boldsymbol{r}, t)$ is the estimated source activity of interest, and $\hat{s}_c(\boldsymbol{r}, t)$ is the estimated background interference plus noise. The problem with the parametric modeling described in the preceding section is that it cannot efficiently take the background interference into account because the Gaussianity assumption may not hold for $\hat{s}_c(\boldsymbol{r}, t)$. The key assumption in the proposed method is that the control measurement that can provide $\boldsymbol{b}_c(t) = \boldsymbol{b}_\xi(t) + \boldsymbol{n}(t)$ is available. In general, this assumption is approximately fulfilled because the pre-stimulus measurement can be considered as a control in many cases. Using this control measurement, the proposed method first derives an empirical distribution of $\hat{s}_c(\boldsymbol{r}, t)$, and with this empirical distribution, the method determines the statistical threshold. The procedures are as follows:

(i) Empirical distribution formation. We calculate $\hat{s}_c(\boldsymbol{r}, t_j)$ by applying the spatial filter to the control measurement $\boldsymbol{b}_c(t_j)$ where t_j is the discrete time point in the control measurement. We then calculate $\hat{F}(x)$, which is the empirical distribution of the modulus of the time course $|\hat{s}_c(\boldsymbol{r}, t_j)|$, such that $\hat{F}(x) = \#\{|\hat{s}_c(\boldsymbol{r}, t_j)| \le x\}/K_c$ where $\#\{|\hat{s}_c(\boldsymbol{r}, t_j)| < x\}$ indicates the number of $|\hat{s}_c(\boldsymbol{r}, t_j)|$ which is less than or equal to x, and K_c is the number of total time points in the control measurement. The implicit assumption here is that the time course reconstruction $|\hat{s}_c(\boldsymbol{r}, t_j)|$ is identically and independently distributed (IID) over the pre-stimulus time period. This procedure is repeated, and the empirical distribution is calculated at all pixel locations. Since $\hat{F}(x)$ is obtained at each pixel location $\boldsymbol{r}$, $\hat{F}(x)$ is rewritten as $\hat{F}(x|\boldsymbol{r})$ below.

(ii) Statistical thresholding without multiple comparisons. Using $\hat{F}(x|\boldsymbol{r})$, we could obtain the statistical threshold at $\boldsymbol{r}$, $\Sigma(\boldsymbol{r})$, such that $\Sigma(\boldsymbol{r}) = \hat{F}^{-1}(1 - \alpha|\boldsymbol{r})$ where α is a level of significance. In practice, the inverse of the empirical distribution can be calculated by first sorting $|\hat{s}_c(\boldsymbol{r}, t_j)|$ in increasing order:

$$|\hat{s}_c(\boldsymbol{r},t_{(1)})| \le |\hat{s}_c(\boldsymbol{r},t_{(2)})| \le \cdots \le |\hat{s}_c(\boldsymbol{r},t_{(K_c)})|, \tag{12}$$

and then by choosing $|\hat{s}_c(\boldsymbol{r}, t_{(q)})|$ as $\Sigma(\boldsymbol{r})$ where $q = [(1 - \alpha)K_c]$ and $|[\bullet]$ indicates the maximum integer that does not exceed the value in parenthesis. However, the statistical threshold obtained in this manner does not take multiple comparisons into consideration, and instead of implementing the above mentioned procedure, the following procedure is performed.

(iii) Statistical thresholding with multiple comparisons. The proposed method uses maximum statistics (Blair and Karniski, 1994; Pantazis et al., 2003) to address the multiple comparison problem. To utilize maximum statistics, we first standardize the empirical distribution of $|\hat{s}_c(\boldsymbol{r}, t_j)|$ by calculating $T(\boldsymbol{r}, t_j)$ such that

$$T(\boldsymbol{r},t_j) = \frac{|\hat{s}_c(\boldsymbol{r},t_j)| - \langle|\hat{s}_c(\boldsymbol{r},t_j)|\rangle_c}{\hat{\sigma}(\boldsymbol{r})} \tag{13}$$

Here,

$$\hat{\sigma}^2(\boldsymbol{r}) = \langle\hat{s}_c(\boldsymbol{r},t_j)^2\rangle_c - \langle|\hat{s}_c(\boldsymbol{r},t_j)|\rangle_c^2,$$

and $\langle\bullet\rangle_c$ indicates the time average over the control period, i.e.,

$$\langle\hat{s}_c(\boldsymbol{r},t_j)^2\rangle_c = \frac{1}{K_c}\sum_{j=1}^{K_c}\hat{s}_c(\boldsymbol{r},t_j)^2 \text{ and}$$

$$\langle|\hat{s}_c(\boldsymbol{r},t_j)|\rangle_c = \frac{1}{K_c}\sum_{j=1}^{K_c}|\hat{s}_c(\boldsymbol{r},t_j)|.$$

We then calculate the maximum T value $T_{\max}(\boldsymbol{r})$ at each pixel location. The maximum T value at the ith pixel location is denoted $T^i_{\max}$, where $i = 1,\ldots, K_N$, and K_N indicates the total number of pixels. We next obtain the empirical distribution of $T^i_{\max}$, $\hat{\mathrm{H}}(x)$, such that $\hat{\mathrm{H}}(x) = \#\{T^i_{\max} \le x\}/K_N$, where $\#\{T^i_{\max} \le x\}$ is the number of $T^i_{\max}$ values which is less than or equal to x. We can then obtain the threshold of the $T^i_{\max}$ value for the α-significance level, $T^{th}_{\max}$, such that $T^{th}_{\max} = \hat{\mathrm{H}}(1 - \alpha)$. The inverse of this empirical distribution can be calculated by first sorting $T^i_{\max}$ in increasing order:

$$T^{(1)}_{\max} < T^{(2)}_{\max} < \cdots < T^{(K_N)}_{\max}, \tag{14}$$

and choose $T^{(p)}_{\max}$ as $T^{th}_{\max}$ where $p = [(1 - \alpha)K_N]$. We finally obtain the statistical threshold for the spatial-filter reconstruction, $\Sigma(\boldsymbol{r})$, by converting $T^{th}_{\max}$ into the source activity value, that is,

$$\sum(\boldsymbol{r}) = T^{th}_{\max}\ \hat{\sigma}(\boldsymbol{r}) + \langle|\hat{s}_c(\boldsymbol{r},t_j)|\rangle_c. \tag{15}$$

We evaluate the statistical significance of the spatial filter outputs by comparing the outputs $|\hat{s}(\boldsymbol{r}, t)|$ with $\sum(\boldsymbol{r})$, and when $|\hat{s}(\boldsymbol{r}, t)| > \sum(\boldsymbol{r})$, the outputs $\hat{s}(\boldsymbol{r}, t)$ are considered to be statistically significant. It should be mentioned that, once we take $T_{\max}(\boldsymbol{r})$ at each pixel location, the distributional properties of the process do not matter anymore, and the assumption that the time course reconstruction is IID is not needed. Instead, this multiple comparison procedure imposes a new assumption that the shape of the control source distribution is the same at each pixel, although its scale may be different.

Numerical experiments

We conducted numerical experiments to show the effectiveness of the proposed statistical thresholding. We use a sensor alignment of the 37-sensor array from the Magnes™ (4D Neuroimaging Inc., San Diego) neuromagnetometer. The source-sensor configuration and the coordinate system are illustrated in Fig. 1(a). The coordinate origin is set at the center of the sensor coil located at the center of the array. The three point-like sources, shown by the small filled circles in this figure, are assumed to be located at (0, −1, −6) cm, (0, l, −6) cm, and (0, 1.6, −7.2) cm on the same plane (x = 0). These locations of the three sources are denoted $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, and $\boldsymbol{r}_3$. The simulated magnetic field is calculated for 400 ms pre-stimulus and 400 ms post-stimulus time windows with a sampling rate of 1 kHz. Here, the nearly orthogonal three time courses shown in Fig. 1(b) are used as the time courses of the three sources.

In these experiments, the pre-stimulus period is considered the control period. The first and second sources are considered the signal sources of interest because they are only active in the post-stimulus (task) period. The third source is considered the control source because it is active both in the pre- and post-stimulus periods. To simulate brain background activity (which is so-called brain noise), spontaneous MEG measured from an awake human subject was added to this computer-generated magnetic field to create simulated magnetic recordings. Here, the spontaneous MEG was measured with a sampling rate of 1 kHz using the same sensor array and averaged over 400 trials. These spontaneous MEG data are shown in the fourth panel (from the top) of Fig. 1(b). The resulting simulated magnetic recordings are shown in the bottom panel of Fig. 1(b).

The eigenspace-projected adaptive spatial filter (Sekihara et al., 2002), mentioned in Typical non-adaptive and adaptive spatial filters, was applied to these simulated recordings. The data between 0 and 400 ms were used for calculating the covariance matrix, and the weight vector of the spatial filter was obtained with this covariance matrix. The dimension of the signal subspace for the eigenspace projection was set at three. The spatial filter was applied to both the pre- and post-stimulus data, and the reconstructed results are shown in Fig. 2. Fig. 2(a) shows the reconstructed time courses at the three source locations, $\hat{s}(\boldsymbol{r}_1, t)$, $\hat{s}(\boldsymbol{r}_2, t)$, and $\hat{s}(\boldsymbol{r}_3, t)$. Three snapshot reconstructions at 220, 265, and 300 ms, $|\hat{s}(\boldsymbol{r}, 220)|$, $|\hat{s}(\boldsymbol{r}, 265)|$, and $|\hat{s}(\boldsymbol{r}, 300)|$ are shown in Fig. 2(b). The reconstruction−averaged over the whole post-stimulus time window between 0 and 400 ms, $\sqrt{\langle \hat{s}(\boldsymbol{r},t)^2 \rangle_{\mathrm{post}}}$ (where $\langle \bullet \rangle_{\mathrm{post}}$ indicates the time average over the whole post-stimulus period)—is shown in Fig. 2(c). The reconstruction region was defined such that $-2.5 < y < 2.5$ cm and $-8 < z < -4$ cm with the pixel interval of 0.5 cm, resulting in the total number of pixels equal to 7373.

To apply the proposed statistical thresholding, we first calculate the modulus of the reconstructed time courses $|\hat{s}(\boldsymbol{r}, t)|$ and derive an empirical null distribution from the pre-stimulus portion of $|\hat{s}(\boldsymbol{r}, t)|$ at each pixel location. The magnitude time courses $|\hat{s}(\boldsymbol{r}_1, t)|$, $|\hat{s}(\boldsymbol{r}_2, t)|$, and $|\hat{s}(\boldsymbol{r}_3, t)|$ are shown in Fig. 3(a). The resultant empirical distributions at $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, and $\boldsymbol{r}_3$, expressed as histograms of the pre-

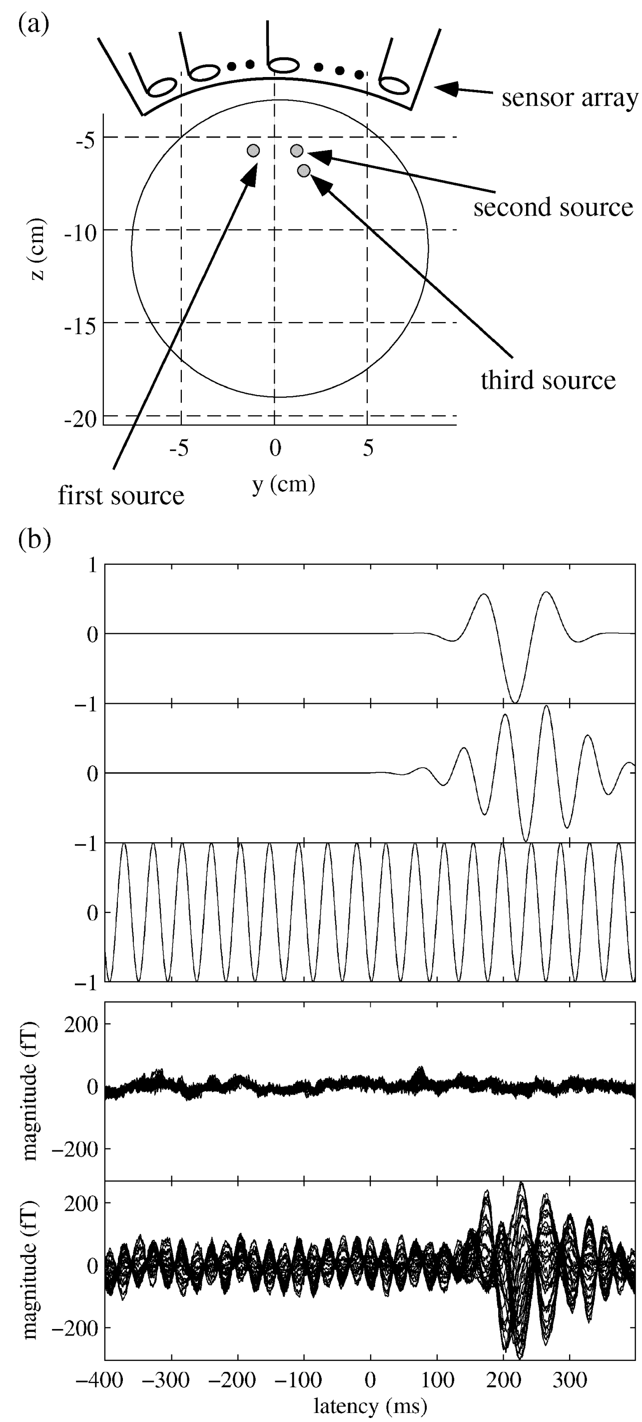

Fig. 1. (a) The coordinate system and source-sensor configuration used in the numerical experiments. The coordinate origin was set at the center of the sensor coil located at the center of the array. The three point-like sources, shown by the small filled circles, were assumed to be located at (0, −1, −6) cm, (0, l, −6) cm, and (0, 1.6, −7.2) cm on the plane of x = 0. The large circle indicates the projection of the sphere used for the forward calculation. (b) The first three panels from top to bottom show the time courses assumed as the time courses of the first, second, and third sources, respectively. The fourth panel shows the spontaneous MEG added to the generated signal magnetic field. The bottom panel shows the simulated magnetic recordings used for the reconstruction experiments.

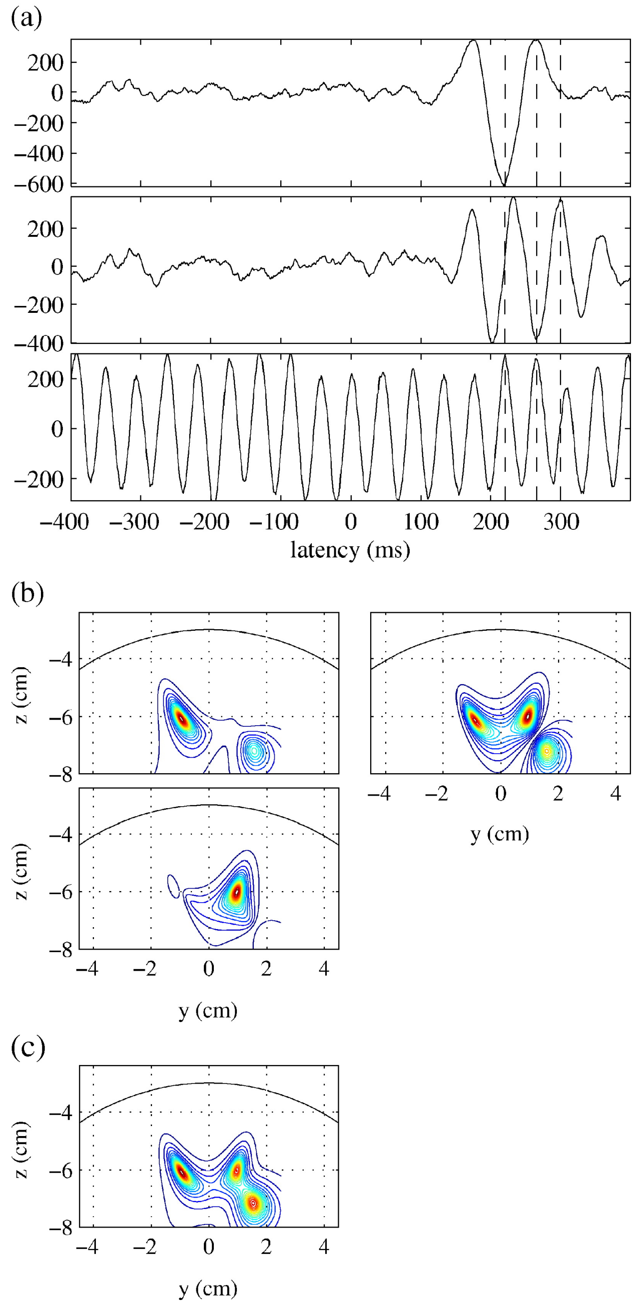

Fig. 2. (a) The reconstructed time courses for the first source $\hat{s}(\boldsymbol{r}_1, t)$ (upper), the second source $\hat{s}(\boldsymbol{r}_2, t)$ (middle), and the third source $\hat{s}(\boldsymbol{r}_3, t)$ (bottom). (b) The snapshot reconstruction at 220 ms $|\hat{s}(\boldsymbol{r}, 220)|$ (upper left), 265 ms $|\hat{s}(\boldsymbol{r}, 265)|$ (upper right), and 300 ms $|\hat{s}(\boldsymbol{r}, 300)|$ (lower left). The time instants at 220, 265, and 300 ms are shown by the three broken vertical lines in panel (a). (c) The reconstruction averaged over the post-stimulus time window, $\sqrt{\langle \hat{s}(r,t)^2 \rangle_{\text{post}}}$.

stimulus values of $|\hat{s}(\boldsymbol{r}_1, t)|$, $|\hat{s}(\boldsymbol{r}_2, t)|$, and $|\hat{s}(\boldsymbol{r}_3, t)|$, are shown in Fig. 3(b). The empirical distributions are then standardized using Eq. (13), and the $T^i_{\max}$, values are derived from the standardized distributions at all pixel locations. The distribution of $T^i_{\max}$ expressed as a histogram is shown in Fig. 4(a). Using this distribution, the value of $T^{th}_{\max}$ is determined to be 3.40 for a significance level α of 0.05. Finally, using Eq. (15), we derive the threshold $\sum(\boldsymbol{r})$, and when $|\hat{s}(\boldsymbol{r}, t)| < \sum(\boldsymbol{r})$, $|\hat{s}(\boldsymbol{r}, t)|$ is set equal to zero. The results of this thresholding with a 5% significance level applied to the reconstruction results in Fig. 2(b) are shown in Fig. 4(b). The thresholded time-averaged reconstruction is shown in Fig. 4(c). The results indicate that the third source, active during both the pre- and post-stimulus periods, is removed from the post-stimulus reconstruction, verifying the effectiveness of the proposed method.

Experiments

We applied the proposed method to auditory-evoked MEG data to test its effectiveness. The auditory-evoked fields were measured using the 275-channel Omega-275™ (VSM MedTech Ltd., Port Coquitlam) whole-cortex biomagnetometer installed at the Biomagnetic Imaging Laboratory, University of California, San Francisco. The auditory stimulus (1-kHz pure tone) was presented to the subject's right ear. The average inter-stimulus interval was 2 s, with the interval randomly varied between 1.75 s and 2.25 s. The sampling frequency was set at 4 kHz, and an on-line filter with a bandwidth from 1 to 2 kHz was used. A total of 400 epochs were measured, and these 400 epochs were averaged to produce the auditory-evoked recordings shown in Fig. 5. Here, although clear P50m and N100m peaks can be observed, we can see that these

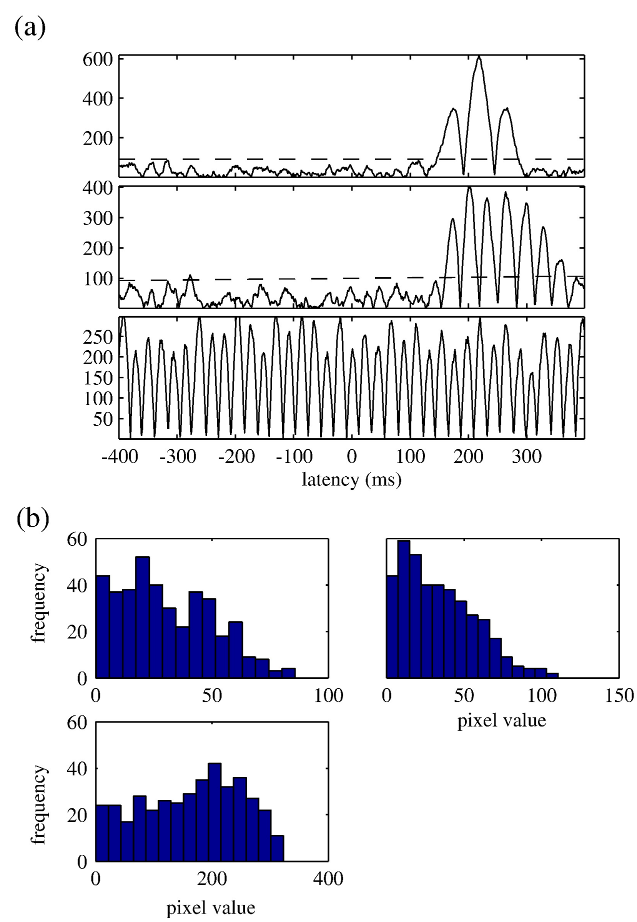

Fig. 3. (a) Reconstructed magnitude time courses, $|\hat{s}(\boldsymbol{r}_1, t)|$, $|\hat{s}(\boldsymbol{r}_2, t)|$, and $|\hat{s}(\boldsymbol{r}_3, t)|$, are shown from top to bottom, respectively. The horizontal broken lines in the upper two panels show the threshold values at $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. (b) Histograms of the pre-stimulus values of $|\hat{s}(\boldsymbol{r}_1, t)|$ (upper left), $|\hat{s}(\boldsymbol{r}_2, t)|$ (upper right), and $|\hat{s}(\boldsymbol{r}_3, t)|$ (lower left).

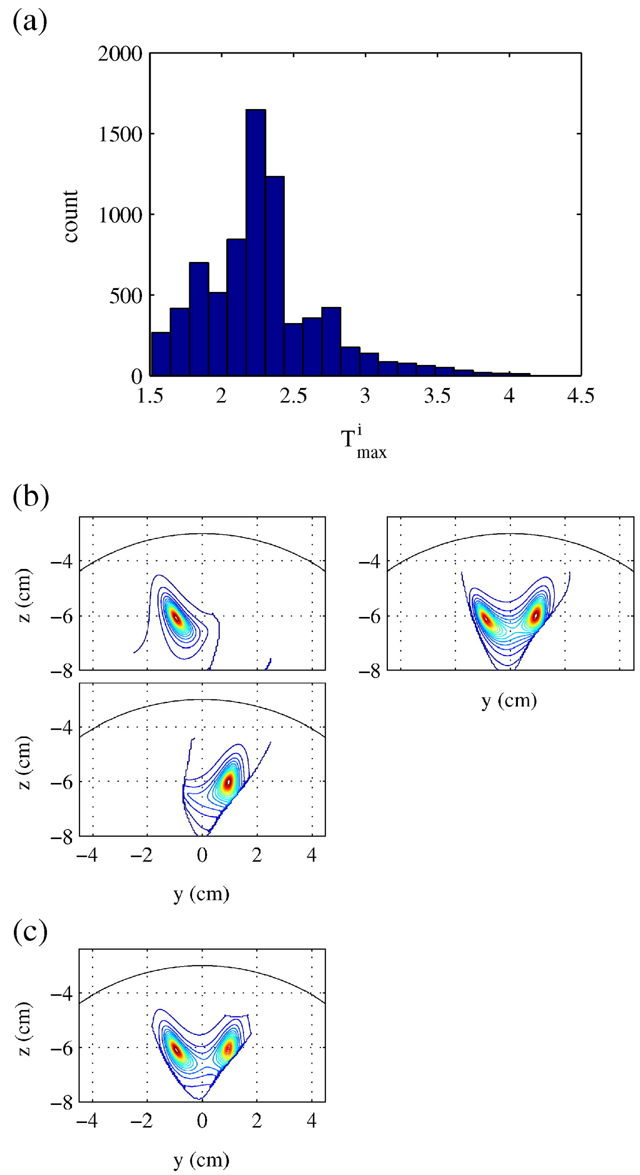

Fig. 4. (a) Histogram of $T^i_{\max}$. The value of $T^{th}_{\max}$ is determined to be 3.40 for a significance level α of 0.05 using this distribution. (b) Thresholded results with $\alpha = 0.05$ for $|\hat{s}(\boldsymbol{r}, 220)|$ (upper left), $|\hat{s}(\boldsymbol{r}, 265)|$ (upper right), and $|\hat{s}(\boldsymbol{r}, 300)|$ (lower left). (c) The thresholded reconstruction for $\sqrt{\langle \hat{s}(\boldsymbol{r},t)^2 \rangle_{\mathrm{post}}}$.

averaged recordings contain a considerable amount of periodic background activity.

The eigenspace-projected adaptive beamformer (Sekihara et al., 2002) was applied to these averaged recordings. The dimension of the signal subspace was set at 2. The data between 0 to 200 ms were used to calculate the covariance matrix **R**, and the weight vector was obtained with this covariance matrix. Two latencies, 44 ms and 86 ms, are selected. One is near the peak of P50m, and the other is near the peak of N100m. These time points are shown in the two vertical broken lines in Fig. 5. The snapshots of the source reconstruction results at these latencies are shown in Figs. 6(a) and (b). Both sets of the results contain a clear localized source in the left temporal lobe probably near the primary auditory area. However, the reconstruction results at 44 ms also contain other diffuse activity. The time-averaged reconstruction obtained from the whole pre-stimulus period (−400 −0 ms), $\sqrt{\langle \hat{s}(\boldsymbol{r},t)^2 \rangle_{\mathrm{pre}}}$, is shown in Fig. 6(c) (where $\langle \bullet \rangle_{\mathrm{pre}}$ indicates the time average over this pre-stimulus period). These results also contain a diffuse source similar to that found in the snapshot at 44 ms, suggesting that these diffuse sources were caused by the periodic background activities observed in the waveforms in Fig. 5.

To apply the proposed statistical thresholding, we use the measurements taken during the pre-stimulus period as a control. We reconstruct the time course $|\hat{s}(\boldsymbol{r}, t_j)|$ of the pre-stimulus period, and the empirical distribution of $|\hat{s}(\boldsymbol{r}, t_j)|$ is calculated at each pixel location. The empirical distributions are then standardized using Eq. (13) to determine $T^i_{\max}$ for the ith pixel, and the same procedure is repeated at all pixel locations to derive the distribution of $T^i_{\max}$. The empirical distribution of $T^i_{\max}$ is shown as a histogram in Fig. 7(a). Here, the number of pixels in the three-dimensional reconstruction grid was 6800. The value of $T^{th}_{\max}$ was determined to be 3.34 using this empirical distribution. Then, using Eq. (15), we derive the threshold $\Sigma(\boldsymbol{r})$. When $|\hat{s}(\boldsymbol{r}, t)| < \Sigma(\boldsymbol{r})$, $|\hat{s}(\boldsymbol{r}, t)|$ is set to zero. The results of this thresholding applied to the snapshot images in Figs. 6(a) and (b) are shown in Figs. 7(b) and (c), respectively. Here, the significance level α was set at 0.05. The diffuse source activities contained in the results in Fig. 6(a) have been removed in Fig. 7(b). These results demonstrate the effectiveness of the proposed statistical thresholding in removing the influence of background source activities in the reconstruction results.

Discussion and conclusion

We have developed a simple and novel nonparametric statistical thresholding procedure for tomographic reconstruction results from

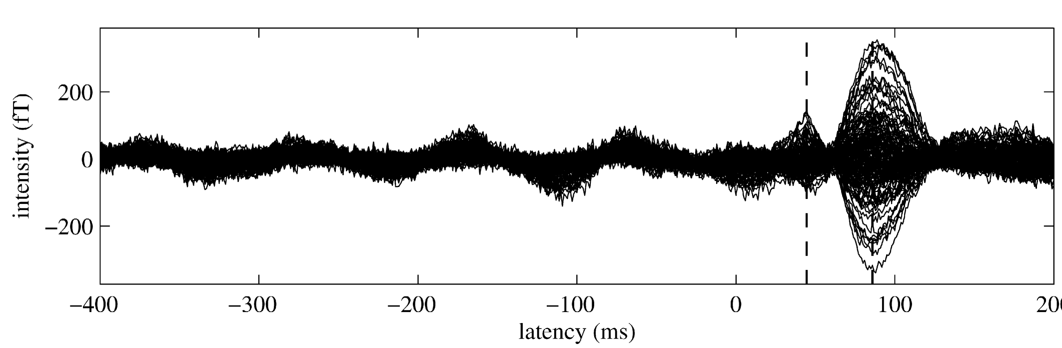

Fig. 5. (a) The 400-epoch-averaged auditory-evoked fields measured using the 275-channel sensor array. Among the 275 sensor recordings, the recordings from 132 sensors covering the subject's left hemisphere are displayed.

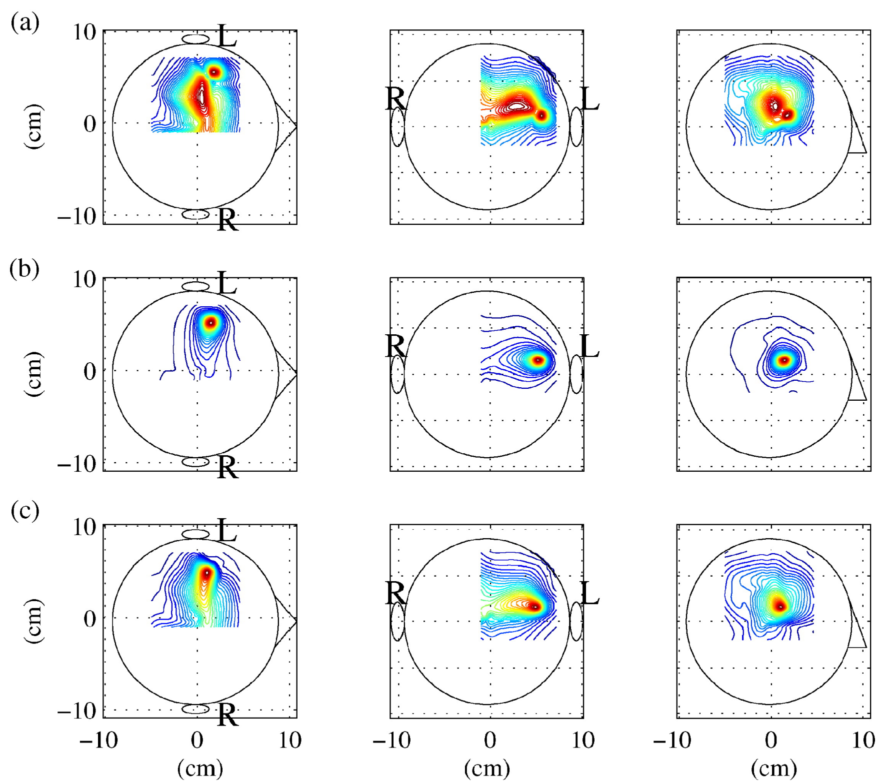

Fig. 6. The maximum-intensity projections of the source reconstruction obtained using the eigenspace-projected adaptive spatial filter. (a) Snapshot at 44 ms of latency, (b) snapshot at 86 ms of latency, and (c) time averaged reconstruction obtained from the whole pre-stimulus period (−400 and 0 ms). The left column shows the maximum intensity projections of the three-dimensional reconstruction onto the axial plane. The middle column shows those onto the coronal plane. The left column shows those onto the sagittal plane. The upper ease letters L and R show the left and the right hemispheres, respectively.

MEG data. Several methods with their aims similar to the one for the proposed method have recently been proposed (Greenblatt and Pflieger, 2004; Pantazis et al., 2003; Singh et al., 2003). Particularly, the methods in (Greenblatt and Pflieger, 2004; Pantazis et al., 2003) can be alternative methods to the one proposed in this paper. These methods use permutation tests to assess the statistical significance of the source reconstruction results. Thus, it not only requires intensive computer time but also requires raw epoch data to be stored. On the contrary, the method proposed in this paper does not use such computer intensive re-sampling methods as the permutation tests or the bootstrap and can thus be implemented with much less computer time. In addition, our method does not require raw epoch data to be stored but uses rather only the averaged data.

The proposed method uses the maximum statistics to address the multiple comparison problems. Without such a multiple comparison procedure, 100α% of pixels may exhibit false-positive activations in a single snapshot image (where α is the level of significance). The multiple comparison procedure, however, reduces this false-positive probability to the level at which only 100α% of snapshot images may contain the false-positive activations. However, since the proposed procedure does not perform the multiple comparisons in the temporal dimension, a reconstructed time course may still contain 100α% of time points that exhibit false activations. This fact can be seen in Fig. 3(a) in which the derived threshold values are indicated by the horizontal broken lines in the upper two panels. Particularly in the middle panel, several time points in the pre-stimulus period exceed the threshold value, resulting in the false-positive activations at these time points. To avoid such false activations, the multiple comparisons should also be performed in the temporal dimension, and for such space–time multiple comparisons, an empirical distribution should be derived not only at each spatial point but also at each time point. Such a method for space–time multiple comparisons is currently under investigation, with the results to be published in the near future.

We pointed out in Proposed nonparametric statistical significance evaluation that we impose the assumption that the time course reconstruction $|\hat{s}_c(\boldsymbol{r}, t_j)|$ is identically and independently distributed (IID) to derive an empirical distribution of $|\hat{s}_c(\boldsymbol{r}, t_j)|$. This assumption is essential for the procedure without multiple comparisons. However, once we take $T_{\max}$ $(\boldsymbol{r})$ at each pixel location, the distributional properties of the process do not matter anymore, and the assumption that the time course reconstruction is IID is not needed for our procedure with multiple comparison. In other words, even if the pre-stimulus process, $\hat{s}_c(\boldsymbol{r}, t_j)$, has strong temporal autocorrelations, we can still compare post-stimulus values to the maximal values of the pre-stimulus process. However, instead of the IID assumption, our multiple comparison procedure imposes a new assumption that the shape of the control source distribution is the same at each pixel, although its scale may be different. This assumption is more attainable than the IID assumption because brain source time courses are known to have significant autocorrelations, but such autocorrelations should be more or less similar at each pixel location.

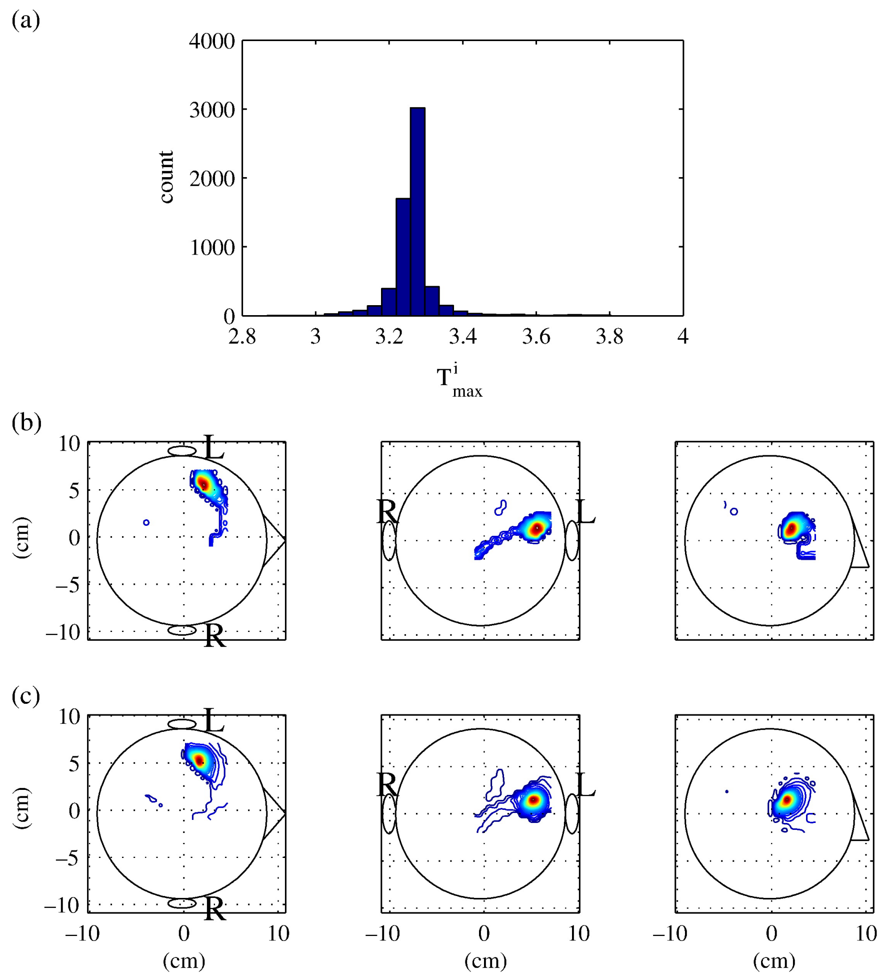

Fig. 7. (a) Histogram of $T^i_{\max}$ from all pixel locations. (b) The results of the proposed statistical thresholding applied to the snapshot shown in Fig. 6(a). (c) The results of the proposed statistical thresholding applied to the snapshot shown in Fig. 6(b).

The other implicit assumption needed for our proposed procedure is that the source activities in the control (pre-stimulus) period are not correlated with those in the task (post-stimulus) period, that is, the control source activities are stationary throughout the task and control periods. This assumption is not exactly hold in such a situation that background (control source) activities are changed before and after the stimulus. Actually, we can often observe that the alpha activity is suppressed by a visual stimulus. In such situations, however, we can assume that the stimulus will also evoke non-phase locked activity that should be accounted for separately from the "pure" control source activities that are unchanged by the stimulus. The proposed method can only remove the influence of the brain sources that are uninfluenced by the stimulus, i.e., "pure" control source activities, and it cannot remove the influence of the stimulus-evoked non-phase locked activity. The proposed method should work for average evoked responses because, in such averaged responses, the contribution of the stimulus-evoked non-phase locked activity is much reduced by averaging, and its influence generally can be neglected.

In Proposed nonparametric statistical significance evaluation, the proposed method is described using the one-tailed test. However, the two-tailed test may be more appropriate when target source activities can vary in the both directions, i.e., when target sources can be deactivated by stimuli. The extension of the proposed method to include the two-tailed test can be performed by first calculating the minimum T value $T_{\min}(\boldsymbol{r})$ at each pixel location where the value of T is calculated using Eq. (13). Denoting the minimum T value at the ith pixel location as $T^i_{\min}$, we next obtain the empirical distribution of $T_{\min}$, $\hat{\mathrm{H}}_{\min}(x)$ such that $\hat{\mathrm{H}}_{\min}(x) = \#\{T^i_{\min} \le x\} / K_N$ and then obtain the threshold of the $T_{\min}$ value for the α-significance level, $T^{th}_{\min}$, such that $T^{th}_{\min} = \hat{\mathrm{H}}^{-1}_{min}(1 - \alpha / 2)$. We can obtain the statistical threshold corresponding to $T^{th}_{\min}$, $\Sigma_{\min}(\boldsymbol{r})$, by using

$$\Sigma_{\min}(\boldsymbol{r}) = T^{th}_{\min}\ \hat{\sigma}(\boldsymbol{r}) + \langle |\hat{s}_c(\boldsymbol{r}, t_j)| \rangle_c. \tag{16}$$

When $|\hat{s}(\boldsymbol{r}, t)| \le \Sigma_{\min}(\boldsymbol{r})$, we can conclude that there is a statistically significant decrease in the source activity at $\boldsymbol{r}$ and time t. The other side of the threshold can be obtained in exactly the same manner as described in Proposed nonparametric statistical significance evaluation except using $T^{th}_{\max} = \hat{\mathrm{H}}^{-1}(1 - \alpha / 2)$.

Although the proposed method has been described as a method for assessing the statistical significance of the source reconstruction results, the method can also be applied to determine the statistical significance of $\boldsymbol{b}_I(t)$ in the sensor recordings $\boldsymbol{b}(t)$. The null hypothesis here is $\boldsymbol{b}_I(t) = 0$ in the task measurements. Let us denote the kth component of the vectors $\boldsymbol{b}(t)$, $\boldsymbol{b}_c(t)$, and $\boldsymbol{b}_I(t)$ as $b^k(t)$, $b^k_c(t)$, and $b^k_I(t)$, respectively. To determine the statistical significance of $b_I^k(t)$, the empirical distribution of $|b_c^k(t_j)|$ is first standardized by calculating

$$T_c^k(t_j) = \frac{|b_c^k\ (t_j)|\ -\ \langle|b_c^k\ (t_j)|\rangle_c}{\hat{\sigma}_b^k}, \tag{17}$$

where

$$\left(\hat{\sigma}_b^k\right)^2 = \langle b_c^k\ (t_j)^2\rangle_c - \langle|b_c^k\ (t_j)|\rangle_c^2,$$

t_j is a discrete time point in the control period, and $\langle \bullet \rangle_c$ again indicates the time average over the control period. We then calculate the maximum T value for the kth channel recording, $T_{\max}^k$, and obtain the empirical distribution of $T_{\max}^k$, $\hat{\mathrm{H}}_b(x)$, such that $\hat{\mathrm{H}}_b(x) = \{T_{\max}^k \le x\} / M^1$. We then obtain the threshold of the $T_{\max}$ value for the α-significance level $T_{\max}^{th}$ such that $T_{\max}^{th} = \hat{\mathrm{H}}_b^{-1}(1 - \alpha)$ and finally obtain the statistical threshold for the kth sensor recording, Σ_b^k, by using

$$\textstyle\sum_b^k = T_{\max}^{th}\ \hat{\sigma}_b^k\ + \langle|b_c^k\ (t_j)|\rangle_c. \tag{18}$$

Therefore, when $b^k(t) > \Sigma_b^k$, the null hypothesis is rejected and we can conclude that a statistically significant $\boldsymbol{b}_I(t)$ exists.

In conclusion, this paper proposes a simple and efficient method of statistical thresholding for MEG spatio-temporal source reconstruction. The method assumes the neural source activities to be quasi-stochastic, and it derives an empirical distribution of non-target sources from the time course reconstruction in the control period. It then derives the value of statistical threshold based on the empirical distribution with the multiple comparisons taken into account. In summary, the method provides an effective means of removing source activities not of interest to the current measurements and of extracting the source activity of interest.

Acknowledgments

This work has been supported by Grants-in-Aid from the Ministry of Education, Science, Culture and Sports in Japan (Cl3680948 and Cl6500296) for KS. This work has been supported in part by the Whitaker Foundation and by the National Institute of Health (P4lRR12553-03 and R0l-DC004855-0lAl) for SN.

References

Baillet, S., Mosher, J.C., Leahy, R.M., 2001. Electromagnetic brain mapping. IEEE Signal Process. Mag. 18, 14–30.

Barnes, G.R., Hillebrand, A., 2003. Statistical flattening of MEG beamformer images. Hum. Brain Mapp. 18, 1–12.

Blair, B.C., Karniski, W., 1994. Distribution-free statistical analysis of surface and volumetric maps. In: Thatcher, R.W. (Ed.), Functional Neuroimaging. Academic Press.

Dale, A.M., Liu, A.K., Fisehl, B.R., Buekner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: combining fMR1 and MEG for high-resolution imaging of cortical activity. Neuron 26, 55–67.

Friston, K.J., Holmes, A.P., Worsley, K.J., Polin, J.-P., Frith, C.D., Frackowiak, R.S.J., 1995. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2, 189–210.

Greenblatt, R.E., Pflieger, M.E., 2004. Randomization-based hypothesis testing from event-related data. Brain Topogr. 16, 225–232.

Gross, J., Kujara, J., Hämäläinen, M., Timmermann, L., Sehnitzler, A., Salmelin, R., 2001. Dynamic imaging of coherent sources: studying neural interactions in the human brain. Proc. Natl. Acad. Sci. 98, 694–699.

Hämäläinen, M.S. Ilmoniemi, R.J., 1984. "Interpreting Measured Magnetic Fields of the Brain: Estimates of Current Distributions," Tech. Rep. TKK-F-A559, Helsinki University of Technology.

Hämäläinen, M., Hari, R., IImoniemi, R.J., Knuutila, J., Lounasmaa, O.V., 1993. Magnetoencephalography—Theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev. Mod. Phys. 65, 413–497.

Hashimoto, I., Kimura, T., Tanosaki, M., Iguehi, Y., Sekihara, K., 2003. Muscle afferent inputs from the hand activate human cerebellum sequentially through parallel and climbing fiber systems. Clin. Neurophysiol. 114, 2107–2117.

Ishii, R., Dziewas, R., Chau, W., Soros, P., Okamoto, H., Gunji, A., Pantev, C., 2003. Current source density distribution of sleep spindles in humans as found by synthetic aperture magnetometry. Neurosci. Lett. 340, 25–28.

Nichols, T.E., Holmes, A.P., 2001. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1–25.

Pantazis, D., Nichols, T.E., Baillet, S., Leahy, R.M., 2003 (July). Spatiotemporal localization of significant activation in MEG using permutation tests. In: Taylor, C.J., Noble, J.A. (Eds.), Proceedings of the 18th Conference on Information Processing in Medical Imaging. Springer Series in Medical Imaging, Ambleside, UK, pp. 512–523.

Pascual-Marqui, R.D., 2002. Standardized low resolution brain electromagnetic tomography (sloreta): technical details. Methods Find. Exp. Clin. Pharmacol. 24, 5–12.

Robinson, S.E., Vrba, J., 1999. Functional neuroimaging by synthetic aperture magnetometry (SAM). In: Yoshimoto, T., et al., (Eds.), Recent Advances in Biomagnetism. Tohoku Univ. Press, Sendai, pp. 302–305.

Sekihara, K., Nagarajan, S.S., 2004. Neuromagnetic source reconstruction and inverse modeling. In: He, B. (Ed.), Modeling and Imaging of Bioelectric Activity—Principles and Applications. Kluwer Academic/Plenum Publishers, pp. 213–250.

Sekihara, K., Seholz, B., 1996. Generalized Wiener estimation of three-dimensional current distribution from biomagnetie measurements. In: Aine, C.I., et al., (Eds.), Biomag 96: Proceedings of the Tenth International Conference on Biomagnetism. Springer-Verlag, New York, pp. 338341.

Sekihara, K., Nagarajan, S.S., Poeppel, D., Marantz, A., Miyashita, Y., 2001. Reconstructing spatio-temporal activities of neural sources using an MEG vector beam-former technique. IEEE Trans. Biomed. Eng. 48, 760–771.

Sekihara, K., Nagarajan, S., Poeppel, D., Marantz, A., Miyashita, Y., 2002. Application of an MEG eigenspace beamformer to reconstructing spatio-temporal activities of neural sources. Hum. Brain Mapp. 15, 199–215.

Sekihara, K., Sahani, M., Nagarajan, S.S., 2005. Location bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. NeuroImage 25, 1056–1067.

Singh, K.D., Barnes, G.R., Hillebrand, A., 2003. Group imaging of task-related changes in cortical synchronisation using nonparametric permutation testing. NeuroImage 19, 1589–1601.

van Veen, B.D., van Drongelen, W., Yuehtman, M., Suzuki, A., 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans. Biomed. Eng. 44, 867–880.

Worsley, K.J., Marrett, S., Neelin, P., Vandal, A.C., Friston, K.J., Evans, A.C., 1996. A unified statistical approach for determining significant signals in images of cerebral activation. Hum. Brain Mapp. 4, 58–73.

[1] $\#\{T_{\max}^k \le x\}$ is the number of $T_{\max}^k$ values which are less than or equal to x.