

Application of Independent Component Analysis with Mixture Density Model to Localize Brain Alpha Activity in fMRI and EEG

Jeong-Won Jeong,¹ Tae-Seong Kim,¹ Sung-Heon Kim,¹ Manbir Singh²

¹ Department of Biomedical Engineering, University of Southern California, Los Angeles, CA

² Departments of Radiology and Biomedical Engineering, University of Southern California, Los Angeles, CA

Received 4 August 2004; accepted 15 August 2004

ABSTRACT: Independent component analysis (ICA) is an approach to solve the blind source separation problem. In the original and extended versions of ICA, nonlinearity functions are fixed to have specific density forms such as super-Gaussian or sub-Gaussian, thereby limiting their performance when sources with different classes of densities are mixed in multichannel data. In this article, we have incorporated a mixture density model such that no assumption about source density would be required. We show that this leads to better source separation due to increased flexibility in handling source-densities with flexible parametric nonlinearity. The algorithm was validated through simulation studies and its performance was compared to other versions of ICA. The modified mixture density ICA was then applied to functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) data to localize independent sources of alpha activity in the human brain. A good spatial correlation was found in the spatial distribution of alpha sources derived independently from fMRI and EEG, suggesting that spontaneous alpha rhythm can be imaged by fMRI using ICA without concurrent acquisition of EEG. © 2004 Wiley Periodicals, Inc. *Int J Imaging Syst Technol*, 14, 170–180, 2004; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/ima.20021

Key words: brain alpha activity; functional magnetic resonance imaging; electroencephalography; independent component analysis

I. INTRODUCTION

Independent component analysis (ICA) is an approach to solve the blind source separation problem within the framework of information maximization (INFOMAX), relying on a negative-entropy maximization, or mutual information minimization in an unsupervised neural network (Bell and Sejnowski, 1995; Hyvärinen et al., 2001). Current ICA algorithms attempt to find a linear weighting matrix that is able to decompose a given data into statistically independent sources whose density forms are assumed as nonlinearity functions at every node of the unsupervised neural network. In the practical applications of these ICAs, it is known that the quality of source separation depends not on the selection of a specific learning method but on the shape of nonlinearity, typically modeled by higher-order

polynomials or hyper-tangent functions under the assumption that all sources are either super-Gaussians (i.e., peak sharper than a Gaussian, positive kurtosis) or sub-Gaussians (i.e., flatter peak than Gaussian, negative kurtosis). Therefore, if the data contains several super-Gaussian and sub-Gaussian sources of different kurtosis, the pre-fixed nonlinearity functions limit the quality of blind source separation. Although there have been several attempts to solve this problem, for example by switching pre-fixed nonlinearity functions or approximating adaptive nonlinearity functions in maximum likelihood estimation, ICA still requires prior information on the nature of the independent components such as the numbers of super- or sub-Gaussian sources and constraint of source variance. Details of these limitations are described in earlier reports (e.g., Hyvärinen et al., 2001; Lee 1998; Lee et al., 1999).

Another limitation of current ICA approaches is that nonlinearity functions in the nodes of unsupervised neural network are fixed only for specific densities. This feature limits the performance of source separation due to its high sensitivity to the signal-to-noise ratio (SNR) (Pearlmutter and Parra, 1997; Xu et al., 1997). To overcome this limitation, a flexible nonlinearity approach has been investigated previously in the INFOMAX algorithm such that no assumption would be required about the source density. In a simulation study, Xu et al. (1997) demonstrated that this adaptive feature of the nonlinearity function reduced the source estimation error by tuning flexible parametric nonlinearity of each channel to the unknown source density without any *a priori* knowledge of the sources such as the number of super- or sub-Gaussian sources and their variances. Heuristic randomization was employed to initialize the shaping parameters of the mixture density model and sequential gradients were utilized to learn the shape of nonlinearity in order to approximate the unknown source density.

However, it was found that an improper random initialization could trap INFOMAX into a local maxima or lead to some non-separation solutions (Xu et al., 1998). This phenomenon becomes more dominant as the data dimension increases and limits the practical application of ICA in modalities such as fMRI or EEG where a relatively large number of sensors are used. In this work, we have implemented a more efficient version of flexible nonlinearity ICA (which we call modified mixture density ICA) to handle the

Correspondence to: Manbir Singh; e-mail: msingh@usc.edu

higher dimensionality of data without requiring any *a priori* knowledge of unknown sources.

The modified mixture density ICA has two features: (1) improved initialization of the shaping parameters to the data probability density using simulated annealing (Press, et al., 1994) and (2) annealing of the learning rate to increase the convergence speed and stability of the solution. The performance of this ICA was compared to conventional ICA methods via computer simulation studies.

In recent years, ICA has been actively applied to the analysis of fMRI and EEG data to decompose the measured signals into statistically independent components. It has been suggested that ICA can isolate the neurological signal components hidden in the measurements and also remove unnecessary artifacts (or noise) efficiently without any specific design of digital filters (Jung et al., 1998; Martin et al., 1998). ICA becomes even more useful in analyzing fMRI and EEG data, when *a priori* models of neural activity are not available and spatiotemporal data are highly contaminated by unknown physiological artifacts and noise. One such example relates to recent attempts to localize the sources of alpha rhythm inside the human brain as described below.

The alpha activity of the brain is a rhythmic pattern of EEG with a characteristic frequency band of 8 to 12 Hz and occurs when the brain is relaxed. It is significantly attenuated with eyes open or mental tasks. Its activity is also intermittent and involuntary. Several investigators have used combined EEG-fMRI measurements to localize alpha activity using conventional model based approaches. These approaches, described below, have relied upon linear correlation analysis (Bandettini, 1993) or SPM (Friston et al., 1996) to detect the activated pixels. The first alpha imaging studies based on near-simultaneous fMRI-EEG measurements were conducted in our laboratory (Patel et al., 1997; Singh et al., 1998), where the concept of using the temporal modulation of the alpha band power to model the predictor (or reference function) was developed. In these studies it was found that the signal, which represented a combination of blood oxygenation level dependent (BOLD) and in-flow effects, decreased in portions of the occipital and parietal lobes. These results are consistent with the first alpha localization studies from EEG alone incorporating a distributed source model (Patel et al., 1999).

Subsequent studies by Goldman et al. (2002) suggested that increased alpha power was correlated with decreased BOLD signal in multiple regions of occipital, superior temporal, inferior frontal, and cingulate cortex, and with increased signal in the thalamus and insula. These results are consistent with Moormann et al. (2003) who reported mostly negative correlation between alpha power and the BOLD signal in the occipital cortex and certain parietal, temporal and frontal regions. However, Laufs et al. (2003) reported a positive correlation between alpha power and the BOLD signal in occipital and mid-cingulate regions and a negative correlation in specific prefrontal and parietal regions by performing multisubject group analysis of continuous and simultaneous EEG/fMRI data. Thus it appears that a negative correlation between alpha power and the BOLD signal is established within the parietal, temporal, and certain prefrontal regions, but at present it is unclear whether there is a positive or negative correlation in certain occipital and frontal regions.

To avoid potential problems of simultaneous fMRI and EEG acquisitions, the study reported here was conducted to investigate the possibility of using fMRI alone, followed by ICA to extract pixels whose time courses conformed to a likely generation of alpha activity. A three-condition fMRI study was conducted with human

subjects where one of the conditions involved closing eyes and relaxing, thus making it a condition likely to generate alpha activity. The other two conditions—eyes open in a lighted room or engaged in a mental arithmetic task—were designed to attenuate alpha activity. EEG data were acquired outside the magnet to verify the generation and suppression of alpha activity from the same subjects during the same three conditions. Alpha rhythms were extracted from the EEG data by using ICA and their sources were localized using a distributed source imaging approach (Khosla, 1996; Patel et al., 1999; Singh et al., 1984). The resulting EEG localizations were spatially correlated to fMRI ICA-determined alpha activity maps suggesting that spontaneous alpha rhythm can be imaged without concurrent acquisition of EEG and fMRI. Details of the methodology and results are described in the following sections.

II. METHOD

A. Background of Independent Component Analysis. Independent component analysis (ICA) is a method for solving the blind source separation problem. Thus the problem is to recover N -independent source signals, $\mathbf{S} = [\mathbf{s}_1; \mathbf{s}_2; \dots; \mathbf{s}_N]$, from N linearly mixed signals, $\mathbf{X} = [\mathbf{x}_1; \mathbf{x}_2; \dots; \mathbf{x}_N]$, where the source \mathbf{s}_i is a zero mean row vector statistically independent from all other sources, and \mathbf{x}_i is a row vector whose elements are linear mixture of all \mathbf{s}_j . Each \mathbf{x}_i is measured in a channel located at different time or space. This relationship between \mathbf{S} and \mathbf{X} can be expressed as a simple matrix form:

$$\mathbf{X} = \mathbf{A}\mathbf{S}, \quad (1)$$

where $\mathbf{A} = [\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_N]$ is unknown full rank matrix and its column vector \mathbf{a}_i defines the mixing weights between the source signals (Martin et al., 1998a).

Given minimal *a priori* knowledge of the nature of the sources and the mixing structure (i.e., sources are statistically independent and linearly mixed in measurement), the task of ICA is to estimate the original source signals, $\mathbf{U} = [\mathbf{u}_1; \mathbf{u}_2; \dots; \mathbf{u}_N]$ from the measurement \mathbf{X} by finding unmixing matrix \mathbf{W} that makes \mathbf{u}_i as statistically independent as possible:

$$\mathbf{U} = \mathbf{W}\mathbf{X}. \quad (2)$$

Because source signals are specified only by a statistical independence, their magnitude and mixing orders cannot be identified. Therefore, we expect that \mathbf{W} is not equal to \mathbf{A}^{-1} but $\mathbf{W}\mathbf{A}$ is equal to $\mathbf{P}\mathbf{C}$. Here \mathbf{P} is an arbitrary permutation matrix whose elements are all zeros except for a 1 in every row and column. \mathbf{C} is an arbitrary scaling matrix whose diagonal elements are only nonzeros. Note that \mathbf{P} and \mathbf{C} are identifiable only if \mathbf{A} is known.

Strict statistical independence yields the general framework of ICA that employs INFOMAX criteria to find the unmixing matrix \mathbf{W} based on an iterative unsupervised neural network implemented with a weighting matrix \mathbf{W} and N nonlinearity functions $\mathbf{y}_i = \mathbf{g}_i(\mathbf{u}_i)$, $1 \leq i \leq N$. The network finds the optimal solution when \mathbf{W} maximizes the joint entropy $\mathbf{Y} = [\mathbf{y}_1; \mathbf{y}_2; \dots; \mathbf{y}_N]$ of $H(\mathbf{Y})$, implying the achievement of minimization of mutual information between estimates of source signals \mathbf{U} .

The gradient of \mathbf{W} to maximize $H(\mathbf{Y})$ is evaluated according to (3) below for each column of \mathbf{X} . The “learning” of \mathbf{W} stops when the norm of gradient converges to zero:

$$\Delta \mathbf{W} = \frac{\partial H(\mathbf{Y})}{\partial \mathbf{W}} \mathbf{W}^T \mathbf{W} = [\mathbf{I} + \rho(\mathbf{U})\mathbf{U}^T] \mathbf{W}, \quad (3)$$

where $\rho(\mathbf{U})$ is called a score function defined as $\rho(\mathbf{U}) = [\rho(\mathbf{u}_1); \rho(\mathbf{u}_2); \dots; \rho(\mathbf{u}_N)]$, $\rho(\mathbf{u}_i) = (\partial p(\mathbf{u}_i)/\partial \mathbf{u}_i)/p(\mathbf{u}_i)$ and $p(\mathbf{u}_i)$ is an estimate of the source probability density function (pdf) modeled as $p(\mathbf{u}_i) = \partial g_i(\mathbf{u}_i)/\partial \mathbf{u}_i$.

Matching of nonlinearity $g_i(\mathbf{u}_i)$ with the cumulative density function (cdf) of the estimated source \mathbf{u}_i is necessary to achieve the global maximum of $H(\mathbf{Y})$. For instance, Bell and Sejnowski (1995) and Lee (1998) showed that if sources \mathbf{S} (mixed in \mathbf{X}) have a fixed form of density such as super-Gaussian (or sub-Gaussian), the pre-fixed nonlinearity function such as a typical logistic sigmoid function (or variant of hyper-tangential function) could be used as a good estimate of cdf of super (or sub-Gaussian) sources to separate these types of sources from the mixture of sources. However, because of the pre-fixed nonlinearity functions, these models do not guarantee to achieve the global solution \mathbf{W} and are limited in separating sources with different densities such as uniform and Gaussian sources.

To overcome these limitations Xu et al. (1997, 1998) showed first that the global solution of INFOMAX could be achieved by adjusting the following model of nonlinearity function to approximate any type of source density.

$$g_i(\mathbf{u}_i) = \sum_{j=1}^J \alpha_{ij} \theta(r_{ij}), \quad (4)$$

where $r_{ij} = b_{ij}(\mathbf{u}_i - a_{ij})$, $\theta(r_{ij}) = 1/(1 + \exp(-r_{ij}))$, $\alpha_{ij} = \exp(\lambda_{ij})/\sum_{j=1}^J \exp(\lambda_{ij})$ for $1 \leq i \leq N$ and $1 \leq j \leq J$.

The nonlinearity function $g_i(\mathbf{u}_i)$ is modeled as a weighted summation of J -logistic sigmoid functions, $\theta(r_{ij})$, that have three different types of shaping parameters: bias a_{ij} , scale b_{ij} , and weight λ_{ij} . A set of parameters $\{\lambda_{ij}, a_{ij}, b_{ij}\}$ is randomly initialized and iteratively updated at the constant rate to approximate the cdf of the source signal \mathbf{s}_i .^{*} Using the flexible nonlinearity function shown in (4) the score function and pdf of an estimate of source \mathbf{u}_i can be written as

$$\rho(\mathbf{u}_i) = \frac{1}{p(\mathbf{u}_i)} \sum_{j=1}^J \alpha_{ij} b_{ij}^2 \theta'(r_{ij}) \quad (5)$$

$$p(\mathbf{u}_i) = \sum_{j=1}^J \alpha_{ij} b_{ij} \theta'(r_{ij}). \quad (6)$$

Here, we focus on the adaptive algorithms of shaping parameters incorporating (1) random initialization and (2) constant learning rates. Their effect on the blind source separation problem will be discussed in the following section.

B. Modified Mixture Density ICA and Its Simulation. The mixture density ICA seeks to adapt the shape of $g_i(\mathbf{u}_i)$ to unknown cdf of sources. Three shaping parameters were randomly initialized and updated by adding a fixed ratio of their gradients to the previous

values. As we can see in (3) these parameters also determine the ongoing unmixing matrix \mathbf{W} . Therefore it is clear that the locality and convergence speed of the solution \mathbf{W} are sensitive to the initialization of shaping parameters. Xu et al. (1998) reported that the mixture density ICA guarantees the convergence of solution \mathbf{W} to the global maxima with high probability if and only if the random initialization leads the initial solution to a nearby region of local maxima.

To guide the initial solution to certain local maxima, we initialize the parameter set to approximate the cdf of the signal measured at each channel. This approximation plays a role in (1) confining the initial range of shaping parameters to the range of the cdf of input signals, thus improving stability by preventing the initial solution from falling in the nonseparable domain and 2) tuning the shaping parameters to have a particular cdf at the initial stage, thus achieving faster convergence. Two main parameters, bias a_{ij} and scale b_{ij} , need to be optimized as

$$\{a_{ij}^*, b_{ij}^*\} = \min_{\{a_{ij}, b_{ij}\}} |g_i(\mathbf{u}_i) - \text{cdf}(\mathbf{x}_i)|^2, \quad (7)$$

where $\text{cdf}(\mathbf{x}_i)$ represents the cdf of \mathbf{x}_i and are approximated by high-order polynomials from \mathbf{x}_i before applying minimization.

In this study, the minimization of (7) is achieved by a simulated annealing method that is able to find the global solutions surrounded by multiple local solutions (Press et al., 1994). Also we add new annealing step to the learning rates η of both \mathbf{W} and $\{\lambda_{ij}, a_{ij}, b_{ij}\}$ to speed the convergence of \mathbf{W} and $\{\lambda_{ij}, a_{ij}, b_{ij}\}$,

$$\eta = \eta_0 \left(1 - \frac{k}{K}\right)^2, \quad (8)$$

where η_0 is the initial value, k is the cumulative number of iteration, and K is the total number of iteration.

In summary, following is the modified mixture ICA algorithm to obtain the solution \mathbf{W} and tune $g_i(\mathbf{u}_i)$ to the unknown source density.

1. Set the number of logistic sigmoid function J . Initialize \mathbf{W} to be identity and weight parameter $\{\lambda_{ij}\}$ randomly in $[0, 1]$. Set η_{w0} , $\eta_{\lambda0}$, η_{a0} , and η_{b0} .
2. Optimize $\{a_{ij}, b_{ij}\}$ according to (7) and use them to initialize $g_i(\mathbf{u}_i)$.
3. For each column of \mathbf{X} compute the gradient of \mathbf{W} and $\{\lambda_{ij}, a_{ij}, b_{ij}\}$ proposed by Xu et al., (1997, 1998). Use them to update \mathbf{W} and $\{\lambda_{ij}, a_{ij}, b_{ij}\}$.
4. If the norm of gradient \mathbf{W} converges to an arbitrary threshold, stop the learning procedure and go to step 5. Otherwise anneal the learning rate η_w , η_λ , η_a , and η_b using (8). Go to step 3.
5. Compute the final estimate of the original sources \mathbf{U} according to (2). Calculate the estimates of cdf and pdf of sources according to (4) and (6).

We carried out an intensive simulation focusing on the effect of flexible nonlinearity on the performance of ICA. This simulation assumed that there exist 10 sources ($N = 10$) measured at 10 channels. Each measurement has 5000 samples ($M = 5000$). Five super-Gaussian and two sub-Gaussian sources (bimodal) were generated using Laplacian density and Pearson mixture density model (Lee, 1998) respectively. Two uniform sources and a Gaussian

^{*} For further details of their adaptive algorithms, see the previous studies listed in Xu et al. (1997, 1998).

source were added to compare the performance of three different versions of ICA. Note that the 10 sources had different kurtosis.

The data matrix \mathbf{X} was acquired by multiplying $\mathbf{S} = [s_1; s_2; \dots; s_{10}]$ by a 10×10 mixing matrix \mathbf{A} in which each element was randomly selected in the range of $[-1 \ 1]$. Three different score functions—one for original ICA (Bell and Sejnowski, 1995), another for extended ICA (Lee et al., 1999), the other for the modified mixture density ICA—were implemented with $\eta_{w0} = 0.0001$ $\eta_{\lambda0} = 0.001$, $\eta_{a0} = 0.01$, $\eta_{b0} = 0.001$, $J = 10$, and $K = 500$. These parameters were also applied to both fMRI and EEG of human subjects.

For all three ICAs the permutation \mathbf{P} and scaling \mathbf{C} were found that make \mathbf{WA} equal to an identity. The performance of each ICA was evaluated in terms of the fractional error e_i between true sources and their estimates.

$$e_i = \frac{\sum_{n=1}^M |s_i(n) - u_i(n)|^2}{\sum_{n=1}^M s_i(n)^2}, \quad (9)$$

where $u_i(n)$ is an estimate of $s_i(n)$ after applying permutation and scaling to the i th row of \mathbf{U} .

We repeated above procedure 50 times with different random sources \mathbf{S} and mixing matrix \mathbf{A} . At each time the fractional error defined in (9) was calculated to compare the separation quality of each ICA.

C. Experimental Design of fMRI and EEG. To obtain experimental fMRI data, normal volunteers were imaged on a 1.5T GE scanner using an echo planar imaging (EPI) sequence with field-of-view or FOV = 24×24 cm², Matrix Size = 64×64 , echo-time or TE = 45 ms, Flip angle = 90°, repetition time or TR = 4 s, 4 oblique contiguous slices, each 1 cm thick, and 125 time-series images per slice. The first five images per slice were disregarded to attain a steady state. Also anatomical images of the four slices and images covering the whole brain were acquired to display the fMRI and EEG results in relation to structure. Functional and structural images were coregistered using the registration method as described by Jeong et al. (2002).

To modulate alpha activity, we designed a 54-s ON-OFF protocol with three conditions, namely, (a) “Relaxation”(ON) where the subject closed eyes and was instructed to relax in a darkened MRI room to induce alpha activity, (b) “Mathematics” (OFF) where the subject kept eyes closed in the darkened room and performed a preassigned arithmetic task to reduce alpha, and (c) “Eyes-open” (OFF) where the subject looked at the lighted MRI room in order to suppress alpha activity. These three conditions were repeated three times in a random order (i.e., “Relaxation”–“Eyes-open”–“Mathematics”–“Eyes-open”–“Mathematics”–“Relaxation”–“Mathematics”–“Relaxation”–“Eyes-open”).

EEG data were obtained outside the magnet before the fMRI experiment using an international standard 10–20 EEG system with the sampling frequency of 256 Hz. The same three-condition experimental paradigm described above was used to acquire EEG data.

D. Localization of Alpha Activity in fMRI Using Data-Driven Method: ICA. Under ICA, each EPI image is placed into subsequent rows, \mathbf{x}_i of the data matrix \mathbf{X} , assuming that each image is a sum of N spatially independent sources whose spatial distributions are defined in subsequent rows, \mathbf{s}_i of the source matrix \mathbf{S} (i.e., spatially independent source maps) and temporal profiles are defined

in subsequent columns of \mathbf{A} (i.e., unique time courses of sources). Here spatial independence indicates that the high values of each map s_i are sparsely focused at specific pixels and rarely overlapped with those of the other s_i . According to this assumption, the time course observed at a given pixel (arbitrary column of \mathbf{X}) can be modeled as a weighted summation of N temporal profiles of \mathbf{A} . The relative weights are represented by the magnitudes of N sources \mathbf{S} at a given voxel. Therefore in ICA, the identification of pixels whose time courses are significantly related to a given experimental protocol can be achieved by the selection of appropriate temporal profiles defined in columns of \mathbf{W}^{-1} and the detection of active values in the maps, \mathbf{u}_i of the selected temporal profiles (Martin et al., 1998a, 1998b).

In this study, the modified mixture density ICA determines the unmixing matrix \mathbf{W} as we proposed in Section II.B. The correlation between the ON-OFF function (“1” for the relaxation and “0” for the other conditions) and the time course associated with \mathbf{u}_i were used to identify the alpha-activity-related component maps, \mathbf{u}_k among 120 spatially independent component maps. The threshold of these correlation coefficients (positive or negative) was set at a given significance level (i.e., critical p value of Spearman rank correlation coefficient) (Blantz, 1992). After selecting the proper \mathbf{u}_k , the elements of \mathbf{u}_k were scaled to Z score vector \mathbf{z}_k and then thresholded at the particular threshold z_{th} (or p value) to detect active elements (Martin et al., 1998b):

$$\mathbf{z}_k = \frac{\mathbf{u}_k - m_k}{\sigma_k} \geq z_{th} \text{ for positive map} \quad (10)$$

$$\mathbf{z}_k = \frac{\mathbf{u}_k - m_k}{\sigma_k} \leq -z_{th} \text{ for negative map,}$$

where m_k and σ_k are the mean and standard deviation of elements in \mathbf{u}_k .

E. Localization of Alpha Activity in EEG Using ICA and Distributed Source Imaging. The distributed source imaging approach (Khosla, 1996; Patel et al., 1999) has been used previously to localize dipoles of evoked potentials (EP) and in magnetoencephalography (MEG) (Singh et al., 1984). The forward problem is formulated in the following way: (1) L dipoles are located at the center of L voxels which are sampled equivalently in a four-sphere head model (Zhou and Van Oosterom, 1992) (i.e., brain, cerebrospinal fluid, skull, and scalp sphere) and (2) at certain time t , the i th electrode potential $v_i(t)$, $1 \leq i \leq N$ (the number of electrodes), is a linear summation of the potentials originating from L dipoles $\vec{d}_l(t)$ and outside noise.

$$\hat{v}_i(t) = \sum_{m=1}^{3L} g_{im} q_m(t) + n_i(t), \quad (11)$$

where $q_{(l-1)+1}(t)$, $q_{(l-1)+2}(t)$, and $q_{(l-1)+3}(t)$ represent three orthogonal components of the l th dipole $\vec{d}_l(t)$ along x, y, and z axis, $1 \leq l \leq L$. $g_{i,3*(l-1)+1}$, $g_{i,3*(l-1)+2}$, and $g_{i,3*(l-1)+3}$ are gains that model the potentials at the i th electrode due to x, y, and z components of $\vec{d}_l(t)$ with unit strength. Gain g_{im} is known as a nonlinear function of the location vector of an individual dipole $\vec{d}_l(t)$, the position vector of the i th electrode, and electrical properties of the head. The gain matrix can be calculated using an analytic

formulation (Mosher et al., 1993). $n_i(t)$ is the noise at the i th electrode.

Based on above forward problem for the electrode potential $v_i(t)$, a maximum entropy (ME) method (Khosla, 1996; Patel et al., 1999) was used to solve the inverse problem, which finds the most probable net dipole strength $|q_i(t)|$ by minimizing the distance between $v_i(t)$ and its estimate $\hat{v}_i(t)$ in the least squares sense. The first application of the distributed source imaging to localize brain alpha activity was proposed by Patel et al., (1999) where the FFT dipole approximation potential $\hat{V}_i(f_a)$ (Lehmann and Michel, 1989) was utilized as the representation of electrode potential $v_i(t)$ in the frequency domain. In this representation, $\hat{V}_i(f_a)$ denotes FFT coefficients of $v_i(t)$ at the harmonic frequency f_a (i.e., f_a represents the alpha activity frequency). The distance of $\hat{V}_i(f_a)$ from the origin, i.e., the average of all $\hat{V}_i(f_a)$ in the cosine-sine diagram (Patel et al., 1999), is called the FFT dipole approximation potential and was used to reveal the alpha activity potential in the frequency domain ME approach. For example, if a single source of alpha activity is oscillating at a specific frequency f_a and phase θ under a noise-free condition, there exist N points of $\hat{V}_i(f_a)$ in the cosine-sine diagram, lying on a straight line with angle θ , since all $\hat{v}_i(t)$, $1 \leq i \leq N$, have an identical phase θ . Therefore each $\hat{v}_i(t)$ can be approximated by the distance of the i th point from the origin. If there are several sources oscillating with different f_a and θ under low SNR, all points do not lie on a straight line since $v_i(t)$ is a mixture of potentials from several sources with different phases. In this case, the projection of each entry to the fitted line was used to calculate the distances of all points from the origin, which degrades the accuracy of the ME method (Patel et al., 1999). To relieve the degradation in the ME method that result from both asynchronous source activity and low SNR, we propose to incorporate ICA into the distributed source imaging approach.

Under ICA, each electrode potential $v_i(t)$ is placed into subsequent rows, \mathbf{x}_i of the data matrix \mathbf{X} , assuming that each potential is a sum of N temporally independent sources whose temporal profiles are defined in subsequent rows, \mathbf{s}_i of the source matrix \mathbf{S} (i.e., temporally independent source profiles) and spatial weights are defined in subsequent columns of \mathbf{A} (i.e., unique spatial distribution of sources). Here temporal independence indicates that the neuronal epochs of each source are activated at different times (i.e., rarely overlapped with those of the other \mathbf{s}_i). According to this assumption, the instant scalp potentials measured at the N electrodes (arbitrary column of \mathbf{X}) can be modeled as weighted summation of N spatial distributions of sources, \mathbf{A} . The relative weights are represented by the magnitudes of N temporal profiles \mathbf{S} at a given time (column of \mathbf{S} corresponding to one of \mathbf{X}). Therefore in ICA, the identification of sources of interest can be done by observing N temporally independent components, \mathbf{u}_p , and their spatial distributions are defined in corresponding columns of \mathbf{W}^{-1} .

To identify \mathbf{u}_i that reflects the temporal profile due to alpha activity, we use the following criterion:

$$\text{SNR}(\mathbf{u}_i) = \frac{\sum_{f=8\text{Hz}}^{12\text{Hz}} |\text{DFT}(\mathbf{u}_i)|^2}{\sum_{f=8\text{Hz}}^{f/2} |\text{DFT}(\mathbf{u}_i)|^2} \geq \text{SNR}_{\text{th}}, \quad (12)$$

where \mathbf{u}_i is the i th row of \mathbf{U} and i represents the index of electrode, $1 \leq i \leq 19$; DFT denotes the discrete time Fourier transform and f_s is the sampling frequency.

Here the SNR represents the ratio of alpha power (i.e., power in the 8–12-Hz band) to background noise power and was used as the basis to detect the significant alpha activity components \mathbf{u}_k .

To display the spatial distribution of alpha power due to each \mathbf{u}_k , we can project \mathbf{u}_k back to the measurement space \mathbf{x}_i through \mathbf{W}_k^{-1} (i.e., the spatial weighting factor for the contribution of \mathbf{u}_k). The projected signal from \mathbf{u}_k to the i th electrode, $\hat{\mathbf{x}}_i$ contains pure alpha activity potential at the i th electrode contributed by only a single alpha component and not any other components such as alpha rhythms with different phases, physiological artifacts, and 60-Hz noise. Therefore FFT dipole approximation potentials $\hat{\mathbf{X}}_i(f_a)$ do not have any deviation in the cosine-sine diagram. This results in the reduction of approximation error. Note that all entries of $\hat{\mathbf{X}}_i(f_a)$ lie on the straight line since all $\hat{\mathbf{X}}_i(f_a)$ have identical phase of a single component \mathbf{u}_k . Because this $\hat{\mathbf{X}}_i(f_a)$ contains one component, its net dipole strengths, $|q_i|$ becomes much more focal than one of raw $\hat{\mathbf{X}}_i(f_a)$. Finally the elements of $|q_i|$ are converted to z-score values and thresholded at the statistical p -threshold in to screen out the active dipoles from spurious dipoles. The active voxels are interpreted as the source map of the component \mathbf{u}_k . The above localization procedure is repeated for all \mathbf{u}_k .

We also calculate the spatial distribution of alpha power, $p_{\text{alpha}}(\hat{\mathbf{x}}_i)$:

$$p_{\text{alpha}}(\hat{\mathbf{x}}_i) = \sum_{f=8\text{Hz}}^{12\text{Hz}} |\text{DFT}(\hat{\mathbf{x}}_i)|^2 \quad (13)$$

and display $p_{\text{alpha}}(\hat{\mathbf{x}}_i)$ on 2D topography.

We conducted a simulation study to show the effects of asynchronous sources and ICA source decomposition on the performance of distributed EEG source imaging. In this simulation we assumed that seven alpha sources were generating 10-Hz sinusoids with different magnitudes and phases. The brain sphere was sampled at 1 cm^3 voxels; the total number of voxels inside the brain sphere was 271. Voxel indexes of the seven source dipoles were 108, 150, 156, 159, 162, 165, and 196 respectively. Based on the above configuration, we calculated the gain matrix to solve the forward and inverse problem in the simulation. Standard conductivity values, radii, and thickness of the four different spheres to model the head as mentioned before (Zhou and van Oosterom, 1992) were adapted from the literature (Mosher et al., 1993). Simulated EEG data were calculated according to (11) (i.e., white Gaussian noise with SNR = 10).

For human studies, we measured the radius and thickness of the four conducting spheres from the sagittal view of the subject's MRI. Three points (nasal, right, and left preauricular points) on the subject's MRI were used as reference points to register the MRI coordinate system to the spherical system (Khosla, 1996). Standard conductivity values were assigned to each sphere (Mosher et al., 1993). Every voxel within the spherical brain region (sized to 1 cm^3) was considered to be a dipole source.

III. RESULTS AND DISCUSSION

A. Simulation Study.

Case 1: Random Mixture. The plots in Figure 1(a) show a typical result of the modified mixture density model for estimating the pdf of unknown sources. Dashed lines represent the true source densities and solid lines their estimates. At each channel the shaping parameters of nonlinearity functions were initialized to approximate the cdf of channel data according to (7) and then estimated through the proposed algorithm described in Section II.B. It is obvious that the resulting density forms of all simulated source signals match their true ones.

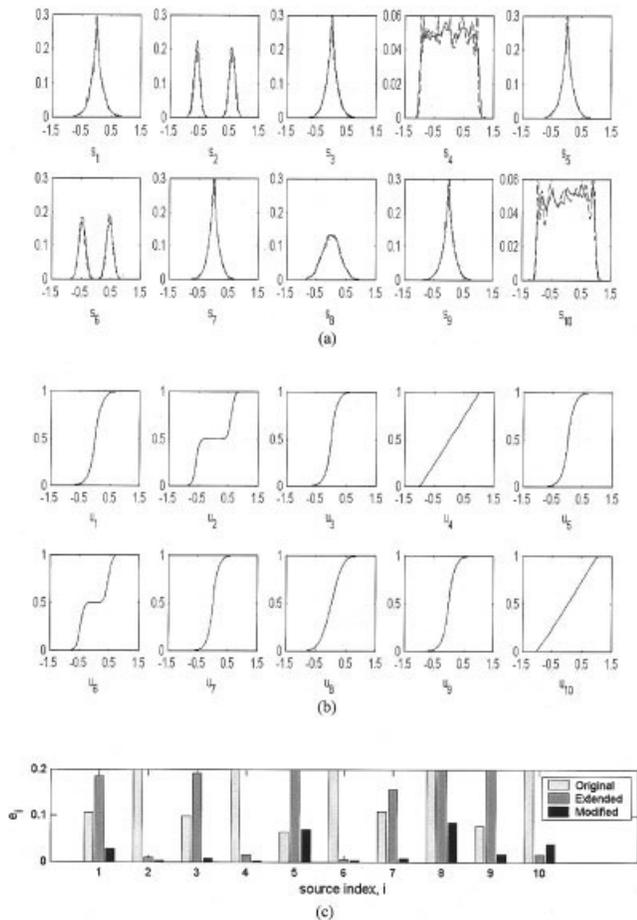


Figure 1. Converged flexible density model and fractional error. (a) Probability density functions of the true sources (solid) and the estimated sources (dashed), super-Gaussian ($\mathbf{s}_1, \mathbf{s}_3, \mathbf{s}_5, \mathbf{s}_7$), sub-Gaussian ($\mathbf{s}_2, \mathbf{s}_4, \mathbf{s}_6, \mathbf{s}_{10}$), and one Gaussian (\mathbf{s}_8). (b) Nonlinearities at all channels, super-Gaussian ($\mathbf{u}_1, \mathbf{u}_3, \mathbf{u}_5, \mathbf{u}_7$), sub-Gaussians ($\mathbf{u}_2, \mathbf{u}_4, \mathbf{u}_6, \mathbf{u}_{10}$), and Gaussian (\mathbf{u}_8). (c) Comparison of fractional errors resulting from three ICAs.

Figure 1(b) shows the flexible nonlinearity $g_i(\mathbf{u}_i)$ of the modified mixture density ICA. It can be seen that the resulting $g_i(\mathbf{u}_i)$ was tuned to approximate the cdf of the unknown source [i.e., each $g_i(\mathbf{u}_i)$ approximates the integration of Figure 1(a)]. This provides good evidence to suggest that the maximization of $H(\mathbf{Y})$ was achieved globally by the solution \mathbf{W} resulting from the shaping parameters of $g_i(\mathbf{u}_i)$. Because all \mathbf{u}_i are independent of each other, all \mathbf{y}_i are statistically independent. Therefore, the maximum of $H(\mathbf{Y})$ is equal to the product of all $H(\mathbf{y}_i)$ (Cover and Thomas, 1991). The maximum of $H(\mathbf{y}_i)$ is achieved if \mathbf{y}_i has a uniform density. This happens at only one condition: $g_i(\mathbf{u}_i) \approx \text{cdf of } s_i$ (Xu et al., 1997).

Figure 1(c) shows the comparison of fractional errors produced by three different ICAs: original, extended, and modified mixture density ICA. Errors were averaged from 50 different \mathbf{S} and \mathbf{A} . The original ICA separated only super-Gaussian group (i.e., $\mathbf{s}_1, \mathbf{s}_3, \mathbf{s}_5$, and \mathbf{s}_7) with a mean error of about 9.14% of total power. The results from the extended ICA also show higher errors in estimating the sub-Gaussians and Gaussian sources than the modified mixture density ICA.

The fractional errors of the modified mixture density ICA became minimal [i.e., about 1.8% of the total power of $s_i(n)$] and

significantly less than those of the original and extended ICA. In its application to blind source separation, the adaptation of nonlinearity could provide the potential to outperform other ICAs with pre-selected and fixed nonlinearity, by allowing more degrees of freedom to the modified mixture density ICA.

Case 2: Mixture of Asynchronous Rhythms. Figure 2(a) shows seven asynchronous alpha sources used to produce simulated EEG data. Note that they are plotted at different colors, with different strengths, and out of phase ($45\text{--}270^\circ$), and partially overlapped with each other. Each plot in Figure 2(b) shows the potentials measured at electrodes of the scalp sphere. The measurement reflects the sum of seven sinusoidal bursts with different phases.

The potentials in the frequency domain were calculated as described in Section II.E. The characteristic frequency of alpha activity was found at the frequency to have the maximal peak in the power spectrum of measurements (i.e., $f_a = 10$ Hz in this simulation). Because sources are asynchronous and white Gaussian noise is

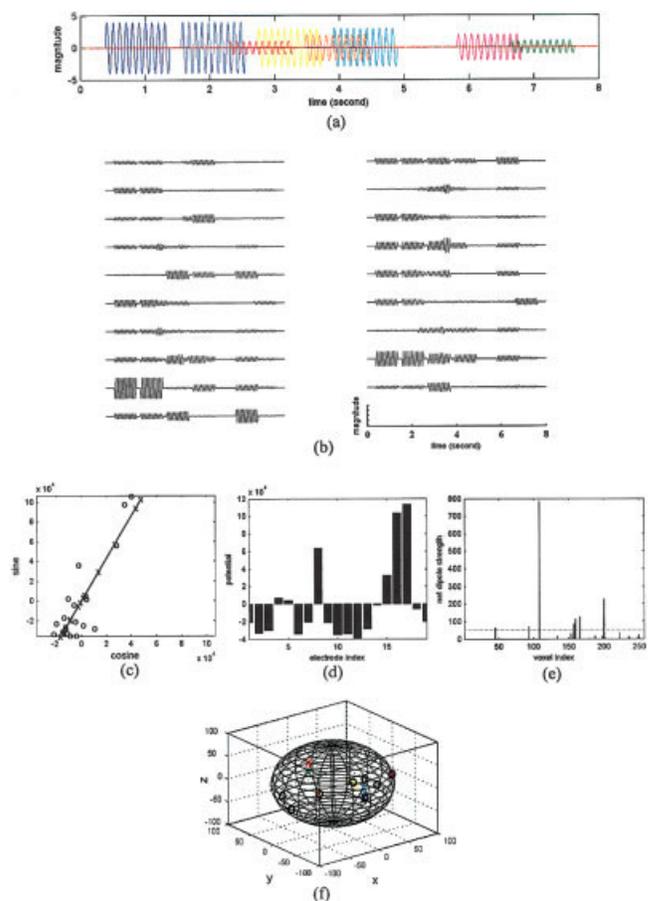


Figure 2. A simulated EEG source localization using the maximum entropy (ME) method in the frequency domain. (a) Seven simulated alpha rhythm sources where each source is coded in a different color. (b) 19-electrode EEG data generated from the forward model. (c) Cosine-sine diagram of (b). (d) FFT dipole approximation potentials of (c). (e) Net dipole strengths obtained from ME method. Note that dotted line indicates the threshold. (f) Locations of active dipoles at p -threshold = 0.01. The different-colored x's indicate the true location of sources corresponding to the colored alpha activity in Figure 2(a) and black o's represent the reconstructed location of these sources.

added, all entries in the cosine-sine diagram are scattered around the least squares fitted line as shown in Figure 2(c). FFT dipole approximation potentials of Figure 2(d) were used as target potentials in the inverse problem of the ME method. The resulting net dipole strengths are shown in Figure 2(e). We applied the p -threshold at 0.01 to reconstruct active dipoles sources from the other dipoles. Figure 2(f) shows the true and reconstructed dipole sources inside the brain sphere. Note that “x” indicates the locations of true dipole sources colored as in Figure 2(a) and “o” represents the reconstructed active dipoles. A total of eight sources at voxel indices of 45, 93, 108, 156, 159, 165, 199, and 200 were detected. Out of these eight sources, only four sources (108, 156, 159, and 165) correspond to true locations.

It was noted that as SNR decreased and phase difference increased the localization error increased proportionally, which results because both factors produce errors in approximating potentials in the cosine-sine diagram. Even though we increased the number of iterations to sharpen the dipole strengths (called the iterative ME method; Patel et al., 1999), the number of reconstructed dipoles corresponding to true sources did not increase any more, which supports our conjecture “both asynchronicity and SNR determine the accuracy of the ME method.”

The modified mixture density ICA was used to isolate the seven alpha components from the simulated data. Alpha dominant components of Figure 3(a) were selected according to (12) with threshold of 1.5. Even though sources were asynchronous, the proposed ICA algorithm could separate all seven sources properly within a mean square error less than 5% of total power. Figure 3(b) shows one example of the projection of alpha dominant component \mathbf{u}_{13} [colored red in Fig. 3(a)] to 19 electrode potentials. It is clear that only the third burst of Figure 2(b) was identified in all electrode potentials.

The cosine-sine diagram from this alpha component was plotted in Figure 3(c). It is clear that there exist no deviation of entries from the fitted line since the projection data originate from a single-phase source, \mathbf{u}_{13} . The FFT dipole approximation potentials of Figure 3(d) were used to localize the net dipole strengths corresponding to this alpha component. The resulting dipole strengths are shown in Figure 3(e) and then thresholded at 0.01 of the p -value. Only one dipole was located at voxel 196, which matches the true location of \mathbf{s}_3 marked as red “x” in Figure 3(f).

The localization of other alpha components is displayed in Figure 3(g). Individual alpha components were in voxel 156 for \mathbf{u}_5 , voxel 59 for \mathbf{u}_8 , voxel 150 for \mathbf{u}_{12} , voxel 196 for \mathbf{u}_{13} , voxel 108 for \mathbf{u}_{16} , voxel 162 for \mathbf{u}_{18} , and voxel 165 for \mathbf{u}_{19} , respectively. They are matched exactly to the voxel index of the original sources \mathbf{s}_1 to \mathbf{s}_7 . Note that “x” and “o” were perfectly matched to each other, implying that direct localization of individual ICA-determined components could enhance the accuracy of the distributed source imaging approach by reducing approximation errors in the representation of electrode potentials.

B. Human Study.

Validation of Experimental Protocol. Figure 4 shows the alpha power spectrum density $p(f)$ of EEG measurement at Fz (1st row) and P4 (2nd row) electrodes during one minute of Relaxation (1st column), Mathematics (2nd column), and Eyes-open (3rd column). Strong alpha activity was clearly observed at the frequency band of around 10–12 Hz during the Relaxation condition, whereas there was relatively low power in the alpha band for either the Mathe-

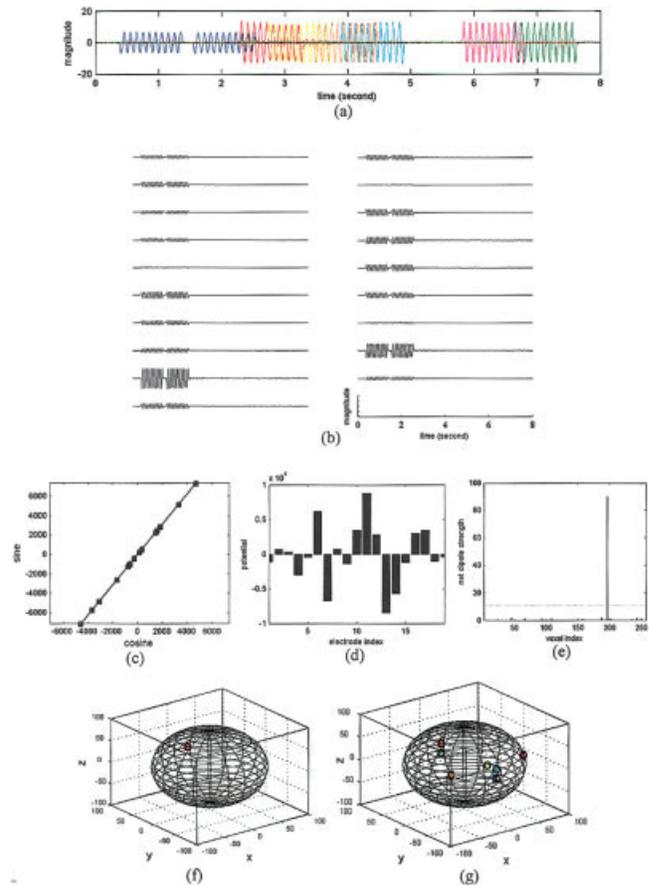


Figure 3. ICA decomposition of simulated data and its localization using ME method. (a) Seven alpha dominant components, each component coded in a different color. (b) Projection of one alpha component: \mathbf{u}_{13} colored red in (a). (c) Cosine-sine diagram from the projection of \mathbf{u}_{13} . (d) FFT dipole approximation potentials of (c). (e) Net strengths of distributed dipolar sources corresponding to (d) obtained from ME method. (f) Location of dipole corresponding to the peak in (e) shown as a black “o”. Red “x” indicates the location of \mathbf{u}_{13} (corresponding to its true location marked as black “o”). (g) The locations of all seven sources corresponding to the seven alpha components (p -threshold = 0.01). The true locations of the other sources are shown by different colors as in (a).

matics or Eyes-open condition. This observation suggests that our three-condition protocol could modulate alpha activity.

Spatially Independent Alpha Components in fMRI. Figure 5(a) shows the slice locations used in this study to cover a large portion of the possible regions of alpha activity. A total of 120 volumes of four slices were rearranged into a matrix whose columns represent the time courses of pixels, decomposed by the modified mixture density ICA.

Among 120 columns of \mathbf{W}^{-1} only two columns, 9th and 14th, showed higher correlation coefficients than the threshold of 0.24 (i.e., $p = 0.01$). The correlation coefficients of the 9th and 14th columns of \mathbf{W}^{-1} with the ON-OFF function were 0.48 and 0.42, respectively. Their corresponding task-related maps, \mathbf{u}_9 and \mathbf{u}_{14} , were thereby chosen to be the spatially independent alpha components. Time courses associated with these maps were shown in Figure 5(c) and (e), respectively.

To detect positive or negative BOLD contrast resulting from alpha activity, pixels were selected according to (10) and color-

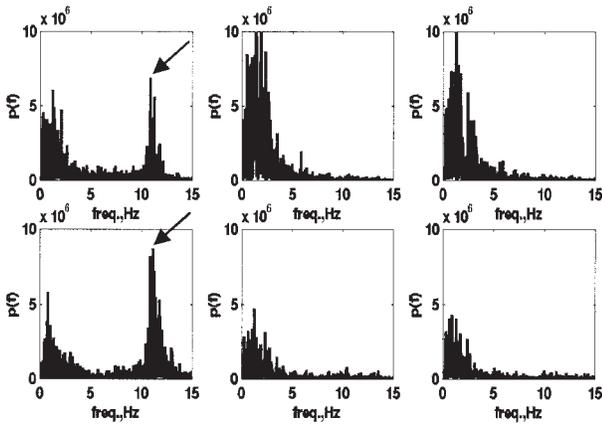


Figure 4. Comparison of alpha power in EEG measured at Fz (frontal, 1st row) and P4 (posterior, 2nd row) electrodes during relaxation (1st column), mathematical task (2nd column) and open eyes conditions (3rd column). The alpha peak (arrow) was clearly seen at 11 Hz during relaxation for this subject, and was absent during the other conditions.

coded red (positive BOLD map) and blue (negative BOLD map) in the anatomical slices. As we can see in the first component shown in Figure 5(b), most activated pixels of the positive map were localized near the precuneus gyrus, right and left angular gyrus in occipital and parietal lobes, consistent with brain regions suggested to be involved in the generation of alpha in previous EEG studies (Lee, 1998). Also the frontal regions around cingulate gyrus, left inferior frontal gyrus, and middle frontal gyrus shown in Figure 5(d) are in accordance with the results of the simultaneous fMRI/EEG studies (Goldman et al., 2002; Laufs et al., 2003; Moormann et al., 2003). Meanwhile in the negative maps shown in Figures 5(b) and (d) we observed relatively sparse activation in left and right frontal lobe, parietal lobe, and motor cortex. These activations were also identified as possible alpha generators in the group analysis of an fMRI/EEG study (Laufs et al., 2003).

Interestingly, major activations of different ICA component maps (i.e., precuneus gyrus in \mathbf{u}_6 and cingulate gyrus in \mathbf{u}_{14}) showed different time courses, suggesting that these areas may be involved in generating asynchronous alpha activity. In these two areas, their averaged time courses increased about $2.45 \pm 1.02\%$ and $2.12 \pm 1.05\%$ respectively, showing slight positive correlation of BOLD signals with the increase of alpha activity. These increased signals during relaxation are consistent with our previous study (Singh et al., 2002) where we observed functional connectivity between cingulate gyrus and the precuneus gyrus. Also, these results are consistent with the study by Greicius et al. (2003) showing that the fMRI activation of the posterior/anterior cingulate cortex increases at the resting epochs but decreases at the working memory epochs. They reported that these precuneus and cingulate gyri were functionally connected and may be critical to the retrieval of episodic memories.

Temporally Independent Alpha Components in EEG. Three components \mathbf{u}_3 , \mathbf{u}_7 , and \mathbf{u}_{12} , showed a relatively higher SNR than the threshold of 1.5. In Figures 6(a)–(c), these three alpha dominant components clearly demonstrate the alpha bursts around 11 Hz. The first topographic map in Figure 6(d) shows the spatial distribution of alpha power from the raw EEG data, and the rest the spatial distribution of each alpha dominant component. The scale in Figure

7(d) indicates the magnitude of alpha power $p_{\alpha}(\tilde{\mathbf{x}}_i)$ using (13). The electrodes are shown as black dots on the display.

As demonstrated in Figure 6(d), the alpha distribution of raw EEG is only focused on the medial occipital lobe; however, the topographies of independent components reveal additional regions of the head which are not obvious in the 2D topography of raw EEG data. The alpha power $p_{\alpha}(\tilde{\mathbf{x}}_i)$ due to \mathbf{u}_3 was highly localized at the Pz and P4 electrodes, which is consistent with the fMRI activation in the medial occipital lobe (i.e., precuneus gyrus), as shown in the second and third slices of Figure 5(b). Interestingly, $p_{\alpha}(\tilde{\mathbf{x}}_i)$ due to \mathbf{u}_7 was concentrated in the O2, O1, and T6 (right end of the parietal lobe) corresponding to activation in the right parietal lobe of the 2nd

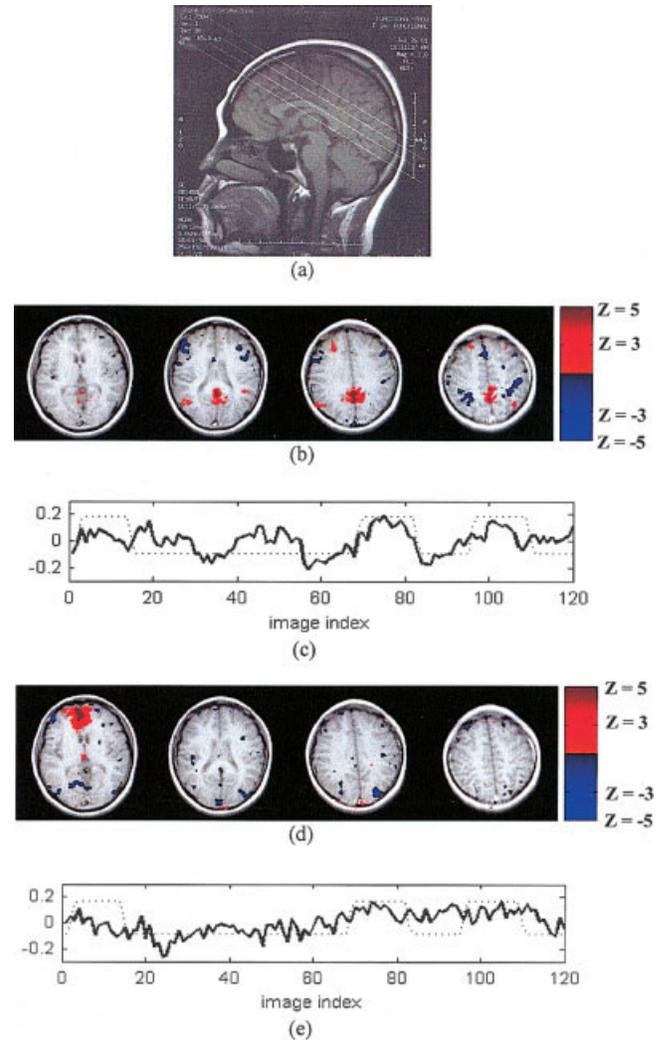


Figure 5. Slice locations and two spatially independent maps with associated time courses highly correlated with experimental paradigm showing increased alpha activity. (a) Location of the four selected oblique slices in the subject's brain. (b) 9th map thresholded at $Z_{th} = 2.57$ ($p = 0.01$). (c) Time course associated with (b) is shown by the solid line. The dotted line denotes ON-OFF block function assuming positive BOLD response during the relaxation. The correlation coefficient between the dotted and solid lines, $cc = 0.48$. (d) 4th map thresholded at $Z_{th} = 2.57$ ($p = 0.01$). (e) Time course associated with (d) and experimental protocol of relaxation, $cc = 0.42$. Red (or blue) of (b) and (d) represents active voxels with higher positive (or negative) Z value, called positive (or negative) alpha map.

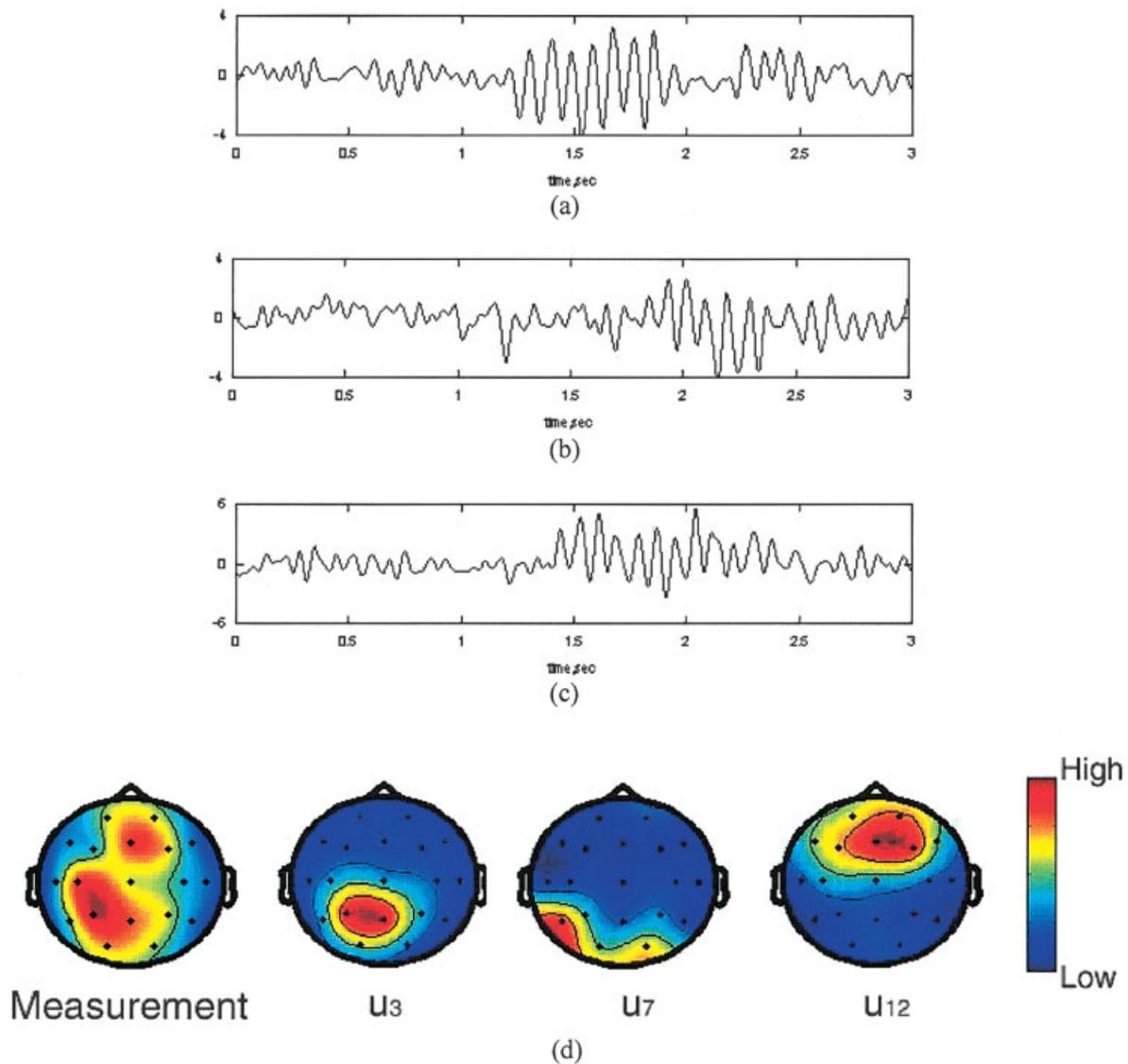


Figure 6. Comparison of measured EEG topography to ICA based topography. (a)–(c) Three alpha dominant components from ICA where (a) is component u_3 , (b) is component u_7 , and (c) is component u_{12} . (d) Spatial distribution of alpha power of the first subject from raw EEG measurement (left), and three alpha dominant components: u_3 , u_7 , and u_{12} (second from the left and toward the right).

and 3rd fMRI slices (i.e., right and left angular gyrus). Also $p_{\text{alpha}}(\vec{x}_i)$ due to u_{12} appears to be localized in Fz (center of frontal lobe), which is consistent with the activation in the 1st slice of the 2nd map shown in Figure 5(d) (i.e., cingulate gyrus).

These three alpha components are in good agreement with the EEG topographical study (Makeig, 2002) showing that at least four alpha activity components (central posterior alpha, left frontocentral alpha, lateral posterior alpha, and right central alpha) contribute to the alpha rhythms. Although source localization was not attempted in the Makeig et al. (2002) study, their topographic locations are consistent with the 3D localizations of the alpha sources obtained by us using ICA.

Localization of Alpha Dominant Components Using EEG Source Localization. Three alpha dominant components— u_3 , u_7 , and u_{12} —were localized separately using the ME method at 0.01 of the p -threshold. We applied this significance level equally to the analysis of fMRI and EEG source localizations (without or with the

mixture density ICA) in order to compare all possible sources at the identical significance level. Figure 7 shows the comparison of the active dipoles identified from raw EEG data (left column) and three alpha dominant components (right column) with the positive and negative fMRI activations shown in Figures 5(b) and (d). In each figure, active EEG dipoles were denoted as green boxes. The red and blue pixels indicate activations corresponding to the positive and negative fMRI maps respectively. The yellow pixels show the overlap of fMRI/EEG activity. The left-column images represent the results of the ME method without ICA and the right-column images those of the ME method with ICA.

We observed that active dipoles from individual alpha dominant components were sparsely distributed in different regions such as the precuneus gyrus of the middle occipital lobe (u_3), near the angular gyrus of the parieto-occipital lobe (u_7), the cingulate gyrus, and middle frontal gyrus of the frontal lobe (u_{12}). Active dipoles around the angular gyrus were 1–3 cm from

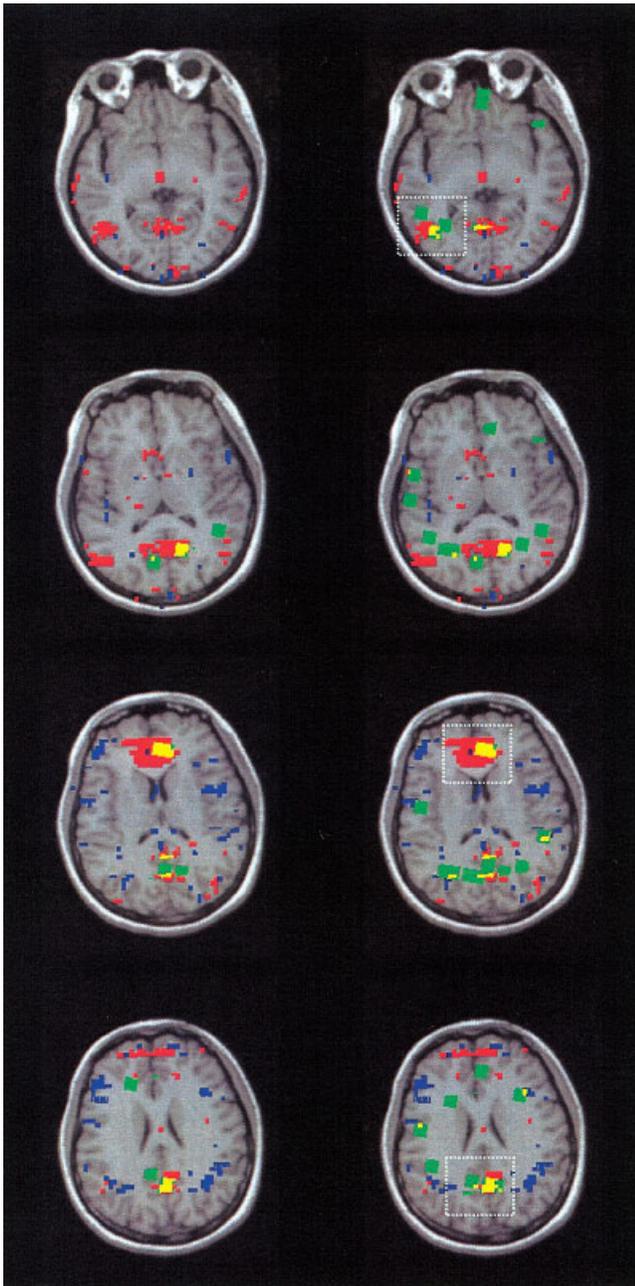


Figure 7. Reconstructed dipole sources in EEG (green squares) and activation in fMRI (positive: red squares; negative: blue squares). Pure EEG data were localized using the ME method without any ICA processing (left column). Each alpha dominant component of Figures 6(a)–(c) was localized separately using the ME method (right column). For both cases, the threshold of net dipole strength was threshold at $\rho = 0.01$. Although overlap between EEG and fMRI is seen in both columns (overlap shown in yellow), overlap is greater when ICA was used. This is clear in the first and last row images as indicated by the white boxes. These two slices were selected as examples because they displayed the greatest density of both EEG and fMRI activations.

the positive activations identified by fMRI as the alpha activity. The locations of these dipoles show good agreement with the positive fMRI activation of this study. See the white boxes in Figure 7.

This study is the first attempt, to our knowledge, to show the spatial consistency of the alpha activity sources in separately measured fMRI and EEG that are fully analyzed by the data-driven ICA approach.

IV. CONCLUSION

We modified the mixture density ICA by incorporating proper initialization and annealing of its shaping parameters such that no assumption about the probability density of sources would be required. With this modification, the convergence of parameters becomes more stable and faster for application to higher dimensional data such as fMRI and EEG. The performance of the modified mixture density ICA algorithm was demonstrated through simulation studies. We show that the ICA formulated in this work is able to separate unknown independent sources with different types of densities with less estimation errors than other ICAs, which have fixed nonlinearity functions depending on *a priori* knowledge of source densities.

The modified mixture density ICA was applied to localize sources of alpha activity in the human brain using fMRI alone and also using EEG alone. It was found that this ICA separates out the alpha activity sources blindly that are highly correlated with the experimental paradigms of both modalities. The ICA-detected alpha activity maps of fMRI are found to be strongly correlated with those of the EEG localization algorithms, suggesting the usefulness of the modified mixture density ICA in analyzing spatiotemporal data like fMRI and EEG.

ACKNOWLEDGMENT

This work is supported in part by grants NIA-NIH P50 AG05142 and NIH-MH RO1 53213.

REFERENCES

- P.A. Bandettini, Processing strategies for time-course data sets in functional MRI of the human brain, *Mag Res Med* 30 (1993), 161–173.
- A. Bell and T. Sejnowski, An Information-maximization approach to blind source separation and blind deconvolution, *Neural Comput* 7 (1995), 1129–1159.
- S. Blantz, *Primer of Biostatistics*, 3rd ed., McGraw-Hill New York, 1992.
- T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley New York, 1991.
- K.J. Friston, A.P. Holmes, K.J. Worsley, J.P. Poline, C.D. Frith, and R.S.J. Frackowiak, Statistical parametric maps in functional imaging: A general linear approach, *Human Brain Mapping* 2 (1996), 185–210.
- R. Goldman, J.M. Stern, J. Engel, and M.S. Cohen, Simultaneous EEG and fMRI of the alpha rhythm, *NeuroReport* 13 (2002), 2487–2492.
- M.D. Greicius, G. Krasnow, A.L. Reiss, and V. Menon, Functional connectivity in the resting brain: A network analysis of the default mode hypothesis”, *Proc Natl Acad Sci USA* 100(1) (2003), 253–258.
- A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley, New York, 2001.
- J. Jeong, T.S. Kim, S. Kim, and M. Singh, Multi-modal MR image registration using mutual information and simulated annealing, *Proc Int Soc Magn Res Med* (2002), 2480.
- T.P. Jung, C. Humphries, T. Lee, S. Makeig, M.J. McKeown, V. Iragui, and T.J. Sejnowski, Extended ICA moves artifacts from electroencephalographic Data, *Adv Neural Inform Process Syst* 10 (1998), 894–900.
- D. Khosla, Functional brain mapping via EEG source localization and its correlation with functional MRI, Ph.D. dissertation, University of Southern California, Los Angeles, 1996.

- D. Khosla, M. Singh, and D. Rice, Three dimensional EEG source imaging via maximum entropy method IEEE Nucl Sci Symp and Med Imaging Conf Rec 151–1519, 1995.
- H. Laufs, A. Kleinschmidt, A. Beyerle, F. Eger, A. Salek-Haddadi, C. Preibisch, and K. Krakow, EEG-correlated fMRI of human alpha activity, *NeuroImage* 19 (2003), 1463–1476.
- T.W. Lee, *Independent Component Analysis: Theory and Applications*, Kluwer Academic, Boston, 1998.
- T.W. Lee, M. Girolami, and T. Sejnowski, Independent component analysis using an extended infomax algorithm for mixed sub-gaussian and super-gaussian sources, *Neural Comput* 11 (1999), 417–441.
- D. Lehmann and C.M. Michel, Intracerebral dipole source of EEG FFT power maps, *Brain Topography*, 2 (1989), 155–164.
- S. Makeig, M. Westerfield, T.P. Jung, S. Enghoff, J. Townsend, E. Courchesne, and T.J. Sejnowski, Dynamic brain sources of visual evoked response, *Science* 295 (2002), 690–694.
- M. Martin, T. Jung, S. Makeig, G. Grown, S. Kindermann, T. Lee, and T. Sejnowski, Spatially independent activity patterns in functional MRI data during the Stroop color-naming task, *Proc Nat Acad Sci USA* 95 1998 (1998a), 803–810.
- M. Martin, S. Makeig, G. Brown, T.P. Jung, S. Kindermann, A. Bell, and T. Sejnowski, Analysis of fMRI data by blind separation into independent spatial components, *Human Brain Mapping* 6 (1998b), 160–188.
- M. Moosmann, R. Ritter, I. Krastel, A. Brink, S. Thees, F. Blankenburg, B. Taskin, H. Obrig, and A. Villringer, Correlates of alpha rhythm in functional magnetic resonance imaging and near infrared spectroscopy, *NeuroImage* 20 (2003), 145–158.
- J.C. Mosher, M.E. Spencer, R.M. Leahy, and P.S. Lewis, Error bounds for EEG and MEG dipole source localization, *Electroenceph Clin Neurophys* 86 (1993), 303–321.
- P. Patel, L. Al-Dayeh, and M. Singh, Localization of alpha-activity by simultaneous fMRI and EEG measurement, *Proc Int Soc Magn Res Med* 3 (1997), 1653.
- P.B. Patel, D. Khosla, and M. Singh, Distributed source imaging of alpha activity using a maximum entropy principle, *Clin Neurophys* 110 (1999), 538–549.
- B.A. Pearlmutter, and L.C. Parra, Maximum likelihood blind source separation: A context-sensitive generation of ICA. In Mozer MC, Jordan MI, Petsche T, Eds. *Advances in Neural Information Processing Systems*, vl. 9, MIT Press, Cambridge, MA, 1997.
- W.H. Press, S.A. Teukosky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C*, 2nd ed., Cambridge University Press, New York, 1994.
- M. Singh, D. Doria, V.W. Henderson, G.C. Huth, and J. Beatty, Reconstruction of images from neuromagnetic fields, *IEEE Trans Nucl Sci* 31 (1984), 585–589.
- M. Singh, J. Jeong, W. Sungkarat, Y. Zhou, and S. Kim, Localization of alpha producing regions in the human brain and their connectivity, *Proc Int Soc Magn Res Med* (2002), 1464.
- M. Singh, P. Patel, and L. Al-Dayeh, FMRI of brain activity during alpha rhythm, *Proc Int Soc Magn Res Med* (1998), 1493.
- L. Xu, C. Cheung, H. Yang, and S. Arami, Independent component analysis by the information- theoretic approach with mixture densities, *Proc 1997 IEEE Int Conf Neural Networks* 3 (1997), 1821–1826.
- L. Xu, C. Cheung, H. Yang, and S. Arami, Learned parametric mixture based on ICA algorithm, *Neurocomputing* 22 (1998), 69–80.
- H. Zhou and A. Van Oosterom, Computation of the potential distribution in a four layer anisotropic concentric spherical volume conductor, *IEEE Trans Biomed Eng* 39(2) (1992), 154–158.