Published in final edited form as:

IEEE Trans Biomed Eng. 2009 November ; 56(11): 2619–2626. doi:10.1109/TBME.2009.2028615.

Source localization of EEG/MEG data by correlating columns of ICA and lead field matrices

Kenneth E. Hild II [Senior Member, IEEE] and Srikantan S. Nagarajan [Senior Member, IEEE]

K. Hild & S. Nagarajan are with the Department of Radiology, University of California at San Francisco, CA, 94143 USA

Abstract

Independent Components Analysis (ICA) has previously been used to denoise EEG/MEG signals before performing neural source localization. Source localization is then performed using a method such as beamforming or dipole fitting. Here we show how ICA can also be used as a source localization method, negating the need for beamforming and dipole fitting. This type of approach is valid whenever an estimate of the forward (mixing) model for all putative source locations is available, which includes EEG and MEG applications. The proposed method consists of estimating the forward model using the laws of physics, estimating a second forward model using ICA, and then correlating the columns of the matrices that represent the two forward models. We show that, when synthetic data are used, the proposed localization method produces a smaller localization error than several alternatives. We also show localization results for real auditory-evoked MEG data.

I. Introduction

Many data sets can be modeled as a linear multiple-input, multiple-output (MIMO) mixture of source signals. The main goal of standard Independent Components Analysis (ICA) methods is to learn a linear MIMO function that, when applied to a mixture of sources (the observations), extracts the original, unmixed sources. The MIMO function that maps the source space to the observation space is commonly referred to as the mixing matrix in the field of ICA and to the forward solution in the field of neuroscience. Likewise, the MIMO function that correctly extracts the sources from the observations is referred to as either the demixing matrix or the inverse solution.

Since its inception researchers have used ICA for a variety of applications. The list of applications includes blind source separation (BSS) [1],[2], blind identification [3],[4],[5], angle of arrival estimation [6], jammer suppression [7], deconvolution [8],[9] feature extraction [10], denoising/artifact rejection [11],[12], genetic analyses [13],[14], and for estimating the number of active sources [15]. Unlike for many of the applications listed previously it is possible to use the laws of physics to estimate the forward solution for electroencephalographic (EEG) and magnetoencephalographic (MEG) applications.

ICA can also be used to estimate the forward solution, but the two estimates have important differences (the differences below were previously pointed out by Knuth [29]). The physics-derived forward solution contains information on the spatial location of all of the putative neural sources, but it is not able to determine which of the putative sources are active. The ICA-derived forward solution depends only on the active sources. It does not, however, contain

any information on the location of these sources. Hence, neither of these approaches can be used, individually, to locate the neural sources.

Localization is commonly performed on EEG and MEG data after applying ICA denoising using, e.g., EEGLAB [17] or its MEG counterpart known as NUTMEG [18]. The ICA denoising is used to extract meaningful components, such as evoked responses, or to remove unwanted interference sources. Localization is then performed on the processed data using techniques such as beamforming [19],[20], dipole fitting [21],[22],[23],[24],[25],[26], or low resolution brain electromagnetic tomography (sLORETA) [27],[28].

In this paper, which extends our previous results [31], we show how the physics-derived and ICA-derived forward solutions can be combined to directly estimate the spatial locations of the active sources, thus negating the need for a separate localization method.

II. Data model

We will assume MEG data throughout, although the ideas can also be applied to EEG data with little modification. The following assumptions are made,

H1: Quasi-static approx. of Maxwell's equations apply

H2: Signals are temporally discrete

H3: Sources are mutually statistically independent

H4: At most one source has a Gaussian distribution

H5: Forward solution is full-column rank

H6: No. of observations $\geq$ no. of sources

Assumption H1 is standard for EEG and MEG data [16], H2 is ubiquitous, and H3-H6 are standard whenever ICA is used [2]. Note, however, that the noise in each sensor represents a unique source in addition to any other sources (thus violating H6). Furthermore, real neural sources are almost certainly distributed (thus violating H5 and H6) and, based on neurophysiological arguments, we expect the time courses of neighboring voxels to be correlated (thus violating H3). Fortunately, as far as ICA is concerned, sources having a relatively small total energy in sensor space can oftentimes be neglected, thus reducing the number of effective sources (blind identification [3],[4],[5] and bias removal [32] may also help alleviate concerns about H6). In addition, multiple focal sources (used to represent a single distributed source) become one effective source (1) in the limit as the maximum pairwise distance between the constituent focal sources approaches zero or (2) in the limit as the minimum pairwise absolute normalized correlation coefficient approaches one. In the former case, it may be possible to accurately localize the effective source when ICA processing is used, but only the mean time course of the constituent focal sources can be recovered. In the latter case, it may be possible to recover the correct time course, but ICA processing necessarily biases the localization. Hence, in cases where an ICA-based method yields meaningful results, it may be that the data are represented well by a relatively few energetic effective sources (where case (1) holds for localization applications and case (2) holds for time course extraction applications) and the effective sources are approximately statistically independent (also, H3 and H4 can be traded for different assumptions commonly used in BSS, e.g., those that require the sources to have unique spectra [38], [41], or to be statistically non-stationary [42]).

Using the quasi-static approximation of Maxwell's equations [16] the (observed) magnetic field at time t can be expressed using the $(N_b \times 1)$ vector, $\mathbf{b}(t)$, as the following volume integral [16],

$$\mathbf{b}(t) = \int_{\Omega} \mathbf{L}(\mathbf{r}) \boldsymbol{\psi}(\mathbf{r}) s(\mathbf{r}, t) d\mathbf{r} + \mathbf{y}(t) + \mathbf{v}(t) \quad (1)$$

where Ω represents the three-dimensional source space, $\mathbf{r}$ is a spatial location within Ω expressed in Cartesian coordinates, $\mathbf{L}(\mathbf{r})$ and $\boldsymbol{\psi}(\mathbf{r})$ are, respectively, the $(N_b \times N_\psi)$ lead field matrix and the $(N_\psi \times 1)$ orientation vector of the neural source at location $\mathbf{r}$, $s(\mathbf{r}, t)$ is the amplitude of the zero-mean neural source (in source space) at location $\mathbf{r}$ and time t , $\mathbf{y}(t)$ is the amplitude of the zero-mean interference sources (in observation space), and $\mathbf{v}(t)$ is the zero-mean noise. If the brain is modeled well as a homogeneous spherical conductor [19], which is a standard assumption for MEG data, then we can use $N_\psi = 2$ since the field outside the head vanishes for radially-oriented dipoles. For the general case $N_\psi = 3$. The neural sources have both an amplitude and an orientation since they can be thought of as infinitesimal current dipoles.

For temporally discrete signals the magnetic field at sample n is assumed to be given by,

$$\mathbf{b}_n = \int_{\Omega} \mathbf{L}(\mathbf{r}) \boldsymbol{\psi}(\mathbf{r}) s_n(\mathbf{r}) d\mathbf{r} + \mathbf{y}_n + \mathbf{v}_n \quad (2)$$

where $\mathbf{b}_n = \mathbf{b}(nT)$, $s_n(\mathbf{r}) = s(\mathbf{r}, nT)$, $\mathbf{y}_n = \mathbf{y}(nT)$, and $\mathbf{v}_n = \mathbf{v}(nT)$, n is the integer-valued sample number, and T is the sampling period. For spatially discrete signals, and under conditions for which we are allowed to interchange the summation and integration, the magnetic field is given by,

$$\begin{aligned} \mathbf{b}_n &= \int_{\Omega} \mathbf{L}(\mathbf{r}) \boldsymbol{\psi}(\mathbf{r}) \sum_{i=1}^{N_s} \delta(\mathbf{r}_i) s_{i,n} d\mathbf{r} + \mathbf{y}_n + \mathbf{v}_n \\ &= \sum_{i=1}^{N_s} \mathbf{L}_i \boldsymbol{\psi}_i s_{i,n} + \mathbf{y}_n + \mathbf{v}_n \\ &= \left[\mathbf{L}_1 \boldsymbol{\psi}_1 \mathbf{L}_2 \boldsymbol{\psi}_2 \cdots \mathbf{L}_{N_s} \boldsymbol{\psi}_{N_s} \right] \mathbf{s}_n + \mathbf{y}_n + \mathbf{v}_n \end{aligned} \quad (3)$$

where $\mathbf{r}_i$ corresponds to the location of the i^{th} neural source, $s_{i,n}$ corresponds to the i^{th} (scalar) element of the $(N_s \times 1)$ vector $\mathbf{s}_n$ (used below), and $\mathbf{L}_i = \mathbf{L}(\mathbf{r}_i)$ and $\boldsymbol{\psi}_i = \boldsymbol{\psi}(\mathbf{r}_i)$ are, respectively, the short hand notations for the lead field matrix and orientation vector associated with the i^{th} neural source. The precise expressions for the lead field matrices, $\mathbf{L}_i$, depend on additional assumptions about the head model. The lead field matrices for several different head models are found in papers authored by Sarvas [16] and Mosher et al. [33].

We define the $(N_b \times N_\psi N_s)$ composite lead field matrix as,

$$\mathbf{L}_c = \begin{bmatrix} \mathbf{L}_1 & \mathbf{L}_2 & \cdots & \mathbf{L}_{N_s} \end{bmatrix} \quad (4)$$

and the $(N_\psi N_s \times N_s)$ orientation matrix as,

$$\boldsymbol{\Psi} = \begin{bmatrix} \psi_1 & 0 & 0 \\ 0 & \psi_2 & 0 \\ 0 & 0 & \cdots & \vdots \\ 0 & 0 & 0 \\ \vdots & \vdots & \psi_{N_s} \end{bmatrix} \quad (5)$$

where $\mathbf{0}$ is a $(N_\psi \times 1)$ vector of zeros. Substituting the product of $\mathbf{L}_C$ and $\mathbf{\Psi}$ from Equations (4) and (5) into Equation (3) produces,

$$\mathbf{b}_n = \mathbf{L}_C \mathbf{\Psi}_{s_n} + \mathbf{y}_n + \mathbf{v}_n \quad (6)$$

We let $\mathbf{F}_s = \mathbf{L}_C$ and $\mathbf{y}_n = \mathbf{F}_x \mathbf{x}_n$, where $\mathbf{F}_s$ is the $(N_b \times N_s)$ forward solution for the neural sources of interest, $\mathbf{F}_x$ is the $(N_b \times N_x)$ forward solution for the interference sources, and $\mathbf{x}_n$ is the $(N_x \times 1)$ vector of interference sources. The final expression for the magnetic field is found by inserting Equations (7) and (8) into Equation (6), which produces,

$$\mathbf{b}_n = \mathbf{F}_s \mathbf{s}_n + \mathbf{F}_x \mathbf{x}_n + \mathbf{v}_n \quad (7)$$

The physical model given by Equation (7) expresses the observations, $\mathbf{b}_n$, as a function of three different types of signals: neural sources, interference sources, and noise. The neural source vector, $\mathbf{s}_n$, is defined as the application-specific sources of interest. The interference source vector, $\mathbf{x}_n$, is defined as the set of signals that are of no interest and are not generated in the sensors. They may originate in the brain, other parts of the body (e.g., heart), or outside the body. The noise is defined as the set of signals that originate in the sensors. Notice that, under H6, Equation (7) is identical to the standard ICA model, with additive noise, which is given

by $\mathbf{b}_n = \mathbf{F} \mathbf{s}_n + \mathbf{v}_n$, where $\mathbf{F} = [\mathbf{F}_s \mathbf{F}_x]$ and $\mathbf{s}_n = [\mathbf{s}_n^T \mathbf{x}_n^T]^T$.

Our goal is to localize each of the neural sources. In our model of the data, given by Equation (7), only the observations, $\mathbf{b}_n$, are known. Anything else we desire must be estimated from the data. We propose to perform localization using two different estimates of $\mathbf{F}_s$, which we hereafter refer to as simply the forward solution (since $\mathbf{F}_x$ is of no interest). We denote the ICA-derived and physics-derived estimates of the forward solution as $\mathbf{A}$ and $\mathbf{F}$, respectively.

III. Proposed method

The proposed source localization method consists of three main steps. The steps are,

1. Estimate ICA-derived forward solution
 - a. Denoise using PFA
 - b. Demix using ICA
 - c. Select neural sources
2. Estimate physics-derived forward solution
 - a. Form grid
 - b. Compute composite lead field matrix
 - c. Estimate orientation matrix
3. Correlate columns of the two forward solutions

A. Estimate ICA-derived forward solution

Many ICA algorithms estimate the inverse solution, as opposed to the forward solution. In this case we need to apply a pseudo-inverse. First, we choose values for N_s and N_x . Then we learn the three matrices that yield the source estimate vector, $\widehat{\mathbf{s}}_n$, and the neural source estimate vector, $\widehat{\mathbf{s}}_n$, which are given by,

$$\begin{aligned}\widehat{\bar{s}}_n &= \mathbf{W}_d \mathbf{W}_p \mathbf{b}_n \\ \widehat{\mathbf{s}}_n &= \bar{\mathbf{I}} \widehat{\bar{s}}_n\end{aligned}\quad (8)$$

where $\mathbf{W}_p$ is a $(N_p \times N_b)$ projection matrix, $\mathbf{W}_d$ is a $(N_p \times N_p)$ demixing matrix, and $\bar{\mathbf{I}}$ is a $(N_s \times N_p)$ selection matrix. Finally, we apply the Moore-Penrose pseudo inverse to the inverse solution, which produces the ICA-derived estimate of the forward solution,

$$\mathbf{A} = \left(\bar{\mathbf{I}} \mathbf{W}_d \mathbf{W}_p \right)^\dagger \quad (9)$$

1) Denoise using PFA—In Step 1a we learn the projection matrix, $\mathbf{W}_p$. The projection matrix is intended to suppress the undesired signals, especially the noise (under the proper conditions the demixing matrix can suppress the interference sources). It also helps to control the complexity of the proposed algorithm by reducing the dimensionality of $\mathbf{b}_n$ prior to learning the demixing matrix. The projection matrix can be found using any desired dimension-reduction method, including methods such as Principal Components Analysis (PCA), Partitioned Factor Analysis, and a supervised method appropriate for brain computer interfaces (BCI) known as the common spatial pattern decomposition [34].

Partitioned Factor Analysis (PFA) denoising [35],[36] is an alternative to PCA. PFA can be used when all or part of the data can be partitioned into two regions: one where the neural sources are known to be active and one where they are known to be inactive. This condition is approximately met when the data are collected using the stimulus-evoked paradigm and then averaged over trials [36]. In this case the active region is assumed to be the post-stimulus window and the inactive region is assumed to be the pre-stimulus window. PFA is designed to suppress both the interference and the noise and it directly extracts the neural sources from the data. Thus, for PFA denoising we (ideally) use $N_p = N_s$.

2) Demix using ICA—In Step 1b we learn the demixing matrix, $\mathbf{W}_d$. The demixing matrix is intended to separate the sources. ICA algorithms typically learn $\mathbf{W}_d$ by making the source estimates, $\widehat{\bar{s}}_n$ as statistically independent as possible. The demixing matrix also enables us to suppress the interference sources through the process of selection.

3) Select neural sources—In Step 1c we learn the selection matrix, $\bar{\mathbf{I}}$. Recall that there is no need for selection if PFA is used, in which case we let $\bar{\mathbf{I}} = \mathbf{I}$. When PFA is not used, we select the subset of $\widehat{\bar{s}}_n$ that are thought to be interference sources. The selection matrix is then formed by taking an identity matrix and removing all rows that are associated with the interference sources. We select the neural sources by taking the N_s source estimates having the largest ratio of post-stimulus energy to pre-stimulus energy.

B. Estimate physics-derived forward solution

The forward solution, $\mathbf{F}_s$, is a function of the location of the N_b sensors and the N_s neural sources. For source localization applications the locations of the sensors are oftentimes known. However, the locations of the neural sources are always, by definition, unknown. To address this shortcoming we form a grid of N_g putative locations and then we use the data to estimate the activity of the putative neural source at each of these voxels. This is the approach taken by beamforming methods [19], which estimate an upper bound of the energy at each voxel. One

consequence of using a grid is that the forward solution, F_s , and its physics-derived estimate, F , differ in size. The former is the $(N_b \times N_s)$ matrix that results from the product of L_C and Ψ , whereas the latter is the $(N_b \times N_g)$ matrix that results from the product of the spatially oversampled versions of L_C and Ψ .

1) Form grid—In Step 2a we build a dense uniform grid of $N_g \gg N_s$ putative locations throughout the brain. The grid is formed over the region corresponding to the brain. The brain region is determined by the co-registration of MEG data with a structural magnetic resonance image (MRI) [37]. More specifically, we place three fiducial coils (one nasion and two opposing preauricular points) on each subject/patient before acquiring their MEG and structural MRI data. We use the NUTMEG suite [18] to load both types of data into memory and then select the brain region with the mouse by clicking on the structural MRI. Once the desired grid resolution is chosen, the NUTMEG program returns r_1 to r_{N_g} , which are the coordinates of the N_g putative source locations.

2) Compute composite lead field matrix—In Step 2b we learn the (spatially oversampled) composite lead field matrix, L_C . This $(N_b \times N_\psi N_g)$ matrix is constructed using Equation (4) with the exception that the neural sources are assumed to be located at the N_g grid points.

3) Estimate orientation matrix—In Step 2c we learn the (spatially oversampled) orientation matrix, Ψ . This $(N_\psi N_g \times N_g)$ matrix is constructed using Equation (5) with the exception that it must use the same set of N_g locations used to create L_C . The $(N_\psi \times 1)$ orientation vector at voxel j , ψ_j , is chosen so that it maximizes a similarity measure between A and the j^{th} column of F . The estimation involves two maximizations. The first is used to find $\psi_{i,j}^*$, which is the optimal value of ψ_j associated with the i^{th} neural source estimate (optimal in the sense of maximizing the similarity metric below). The second is used to find the optimum value taken over all N_s neural source estimates.

The parameter $\psi_{i,j}^*$ is given by,

$$\begin{aligned} \psi_{i,j}^* &= \underset{\psi_j}{\operatorname{argmax}} \frac{|A_i^T F_j|}{(A_i^T A_i F_j^T F_j)^{1/2}} \\ &= \underset{\psi_j}{\operatorname{argmax}} \frac{\psi_j^T L_j^T A_i A_i^T L_j \psi_j}{\psi_j^T L_j^T L_j \psi_j} \end{aligned} \quad (10)$$

where A_i is the i^{th} column of A , F_j is the j^{th} column of F , and where we use the fact that the square root operator monotonically increases over the domain, the maximization is independent of the positive scalar quantity $A_i^T A_i$ and $F_j^T F_j$ equals $L_j^T L_j$. It is obvious that the optimal solution is independent of the norm of ψ_j . Hence, we add the constraint that ψ_j must have unit norm.

Notice that both $L_j^T A_i A_i^T L_j$ and $L_j^T L_j$ are symmetric and positive semi-definite matrices. Hence, the vector ψ_j that maximizes (15) is given by the eigenvector associated with the largest eigenvalue of the generalized eigenvalue decomposition (GED) of $L_j^T A_i A_i^T L_j$ with respect to $L_j^T L_j$. The overall optimum orientation vector at the j^{th} voxel is found by jointly optimizing over all possible orientation vectors and all possible neural sources. This is accomplished by replacing the subscript i in Equation (10) with $f(j)$, which is defined by,

$$\begin{aligned}
f(j) &= \underset{i}{\operatorname{argmax}} \frac{|A_i^T F_{i,j}^*|}{(A_i^T A_i F_{i,j}^{*T} F_{i,j}^*)^{1/2}} \\
&= \underset{i}{\operatorname{argmax}} \frac{|A_i^T L_j \psi_{i,j}^*|}{(A_i^T A_i \psi_{i,j}^{*T} L_j^T L_j \psi_{i,j}^*)^{1/2}}
\end{aligned} \tag{11}$$

where $F_{i,j}^* = L_j \psi_{i,j}^*$. We denote the overall optimum orientation vector as $\psi_{f(j),j}^*$.

C. Correlate columns of the two forward solutions

If assumptions H1 and H2 are valid, the orientation vector at each (true) neural source location is estimated accurately, and if the true locations of the neural sources happen to fall on the chosen grid, which is guaranteed for uniform sampling in the limit as $N_g \rightarrow \infty$, then (by definition) each of the N_s columns of $\mathbf{F}$ that correspond to a neural source is identical to one of the columns in $\mathbf{F}_s$.

If, in addition, assumptions H3 to H6 are valid, the userdefined value of N_s equals the true number of neural sources, the combination of denoising and demixing successfully removes all interference and noise, and $\mathbf{A}$ is such that the neural source estimates are mutually statistically independent then each of the N_s columns of $\mathbf{A}$ equals (to within a multiplicative scalar) one of the columns in $\mathbf{F}_s$.

The consequence of the two preceding paragraphs is that each of the columns of $\mathbf{A}$ also equals (to within a multiplicative scalar) one of the columns in $\mathbf{F}$. Although each column of $\mathbf{F}$ corresponds to a known location in source space (under the above conditions), we do not know which of the columns correspond to (active) neural sources. Likewise, each column of $\mathbf{A}$ corresponds to an active neural source, but $\mathbf{A}$ does not provide information on the location of the associated sources. In the proposed approach we combine the information contained in $\mathbf{F}$ and $\mathbf{A}$ by comparing their columns. More specifically, we use the following similarity metric to express the confidence level that a neural source exists at voxel j ,

$$\begin{aligned}
\rho_j &= \frac{|A_{f(j)}^T F_{f(j),j}^*|}{(A_{f(j)}^T A_{f(j)} F_{f(j),j}^{*T} F_{f(j),j}^*)^{1/2}} \\
&= \frac{|A_{f(j)}^T L_j \psi_{f(j),j}^*|}{(A_{f(j)}^T A_{f(j)} \psi_{f(j),j}^{*T} L_j^T L_j \psi_{f(j),j}^*)^{1/2}}
\end{aligned} \tag{12}$$

where ρ_j is the absolute value of the normalized correlation coefficient (except that the means of the columns are not removed), which is guaranteed to take values between 0 and 1, and Equation (11) is used to determine $f(j)$, which is the column of $\mathbf{A}$ that is most similar to the j^{th} column of $\mathbf{F}$. The means are not removed from the columns in the similarity metric because they represent parameters of a system, as opposed to a signal, in which case differences in the mean are significant. Notice that Equations (15) and (16) also use the same similarity metric given by Equation (12).

Due to the well-known gain and permutation indeterminacies of ICA [2], the ICA-derived estimate of the forward solution, under the ideal conditions listed above, is related to the true forward solution using $\mathbf{F}_s = \mathbf{DPA}$, where $\mathbf{D}$ is a diagonal matrix and $\mathbf{P}$ is a permutation matrix. Because of the permutation ambiguity of ICA we do not know which column of $\mathbf{F}$ is associated with a given column of $\mathbf{A}$. For this reason we compare all columns of $\mathbf{F}$ with all columns of $\mathbf{A}$. Likewise, because of the gain ambiguity of ICA we use a similarity metric that is invariant to multiplicative gains.

IV. Results

A. Synthetic Data

For the synthetic data we use a total of 500 different evoked-response data sets, which corresponds to 50 Monte Carlo trials for each of 2 experiments and 5 different values of the input signal-to-interference ratio (SIR), which is defined by,

$$\text{SIR} = 10 \log_{10} \left(\frac{\sum_{i=1}^{N_b} \sum_{n=0}^{N_{\text{post}}-1} \left(\sum_{k=1}^K \mathbf{F}_s \mathbf{s}_{n,k} \right)_i^2}{\sum_{i=1}^{N_b} \sum_{n=0}^{N_{\text{post}}-1} \left(\sum_{k=1}^K \mathbf{F}_x \mathbf{x}_{n,k} \right)_i^2} \right) \text{ (dB)} \quad (13)$$

where the stimulus onset is $n=0$, $\mathbf{s}_{n,k}$ and $\mathbf{x}_{n,k}$ are $(N_s \times 1)$ vectors, $\left(\sum_k \mathbf{F}_s \mathbf{s}_{n,k} \right)_i^2$ is the square of the (scalar) i^{th} element of the sum, and $\left(\sum_k \mathbf{F}_x \mathbf{x}_{n,k} \right)_i^2$ is similarly defined. Spatially and temporally white Gaussian noise is added to each sensor. The power of the isotropic noise is chosen so that the input signal-to-noise ratio (SNR) is -5 dB. Each dataset is 1000 samples in length, 5/8 of which corresponds to the pre-stimulus period and 3/8 to the post-stimulus period, and the sampling frequency is 1 kHz. Each neural source has a damped sinusoidal time course having a random frequency between 5–20 Hz, whereas each (neural) interference source has an un-damped sinusoidal time course. Damping is performed such that the energy of the neural sources is concentrated within the post-stimulus period. Determination of the lead field matrices requires knowledge of the spatial locations of the sensors. For this we use the true spatial locations of the sensors corresponding to an in-house 275-channel whole head MEG system (Omega 2000, VSM MedTech Inc., Port Coquitlam, Canada). To simplify the analysis all of the sources of the synthetic data are placed within the sagittal plane that passes through the center of the head. The voxel size is 3.7 mm per side and $N_g = 1600$. The number of sensors is $N_b = 275$, the number of neural sources is $N_s = 2$, and the number of interference sources is $N_x = 3$. We assume that the true values of N_s and N_x are known.

The performance is measured using the localization error, which is the mean distance between the true neural source locations and the estimated locations. The estimated locations correspond to the tallest two peaks of the tomographic map. Since there are 2 neural sources, there are two different ways of associating the true source locations with the estimated source locations. We use the more meaningful association, which is the one that has a smaller mean distance.

We compare the performance of the proposed column correlation method, 3 scalar minimum variance beamformers (operating on the raw data and operating on denoised data using PCA and PFA denoising), and 2 methods that use ICA denoising prior to localization (sLORETA and dipole fits). The proposed column correlation method uses PFA denoising and the AR-MOG ICA algorithm with $K_g = 0$ and $K_Q = 3$ [38]. The AR-MOG algorithm assumes each source distribution is a mixture of Gaussians, the parameters of which are learned from the data. Hence, it is appropriate for EEG and MEG sources, which are thought to have multi-modal distributions [40]. For the beamforming methods we use the denoised data covariance matrix whenever denoising is used. For all 6 methods we use Bayesian regularization for inverting ill-conditioned matrices, where a matrix is defined to be ill-conditioned when the ratio of the maximum eigenvalue to the minimum eigenvalue is greater than $1e12$.

First we report the computational complexities of the 6 methods, which are measured in terms of mean completion time per dataset. The mean completion times are 3.0, 3.1, 18.7, 19.3, 40.6,

and 63.0 seconds for the PCA/Beamformer, Raw Data/Beamformer, PFA/Beamformer, PFA/Correlate-Columns, ICA/Dipole Fit, and ICA/sLORETA, respectively.

Figures 1 and 2 pertain to the first experiment. Figure 1 shows contour plots for 4 of the methods, where the y-axis represents left-right position and the z-axis represents vertical position. For this example, the Raw Data/Beamformer has peaks at all 5 (neural and interference) sources (the peaks at the neural sources are barely visible), the PCA/Beamformer map has a single peak occurring between 2 of the interference sources, whereas the PFA/Beamformer and PFA/Column Correlation methods have peaks near both neural sources.

Figure 2 shows the mean localization error for the first experiment as a function of the (input) SIR. The ICA/Dipole Fits method, not shown in order to reduce clutter, performs slightly better than PCA/Beamforming for low SIR and slightly worse than PCA/Beamforming for high SIR. The ICA/Dipole Fits and PCA/Beamformer methods consistently perform the worst, PFA/Column Correlation has the best performance for most values of SIR, and the Raw Data/Beamformer performs poorly for low SIR and performs the best (as expected) for extremely high SIR.

Figures 3 and 4 pertain to the second experiment. In this experiment we replace one of the focal neural sources with a distributed source. The single distributed source spans a $1 \text{ cm} \times 1 \text{ cm}$ area and consists of 9 uniformly-placed focal sources, each of which has an independently-drawn time course and orientation. Each of the methods continues to assume that there are only 2 sources of interest, although there are actually 10. Figure 3 shows the contour plots of 4 of the methods. In this example, the Raw Data/Beamformer has a single large peak at one of the interference sources and the PCA/Beamformer has peaks at 2 of the interference sources, whereas PFA/Beamformer and PFA/Column Correlation have peaks near both of the neural sources and 1 or 2 extraneous peaks.

Figure 4 shows the mean localization error for the second experiment. The ICA/Dipole Fits method (not shown) performs the same as ICA/sLORETA for low SIR and slightly worse than PCA/Beamforming for high SIR. The ICA/Dipole Fits, ICA/sLORETA, and PCA/Beamformer methods consistently perform the worst and the PFA/Column Correlation consistently performs the best.

B. Real Data

Next we present results from real auditory-evoked data from a male subject in his mid 20's. The stimulus is a 1 kHz tone of 400 ms duration, which is presented binaurally once every 1.35 sec (± 100 ms jitter). The sample rate is 4 kHz, the number of trials is 247, and each trial lasts for 800 ms. We apply a 4 Hz high pass filter to the data in addition to the hardware anti-alias filter. The structural MRI data are collected using a 1.5 T Signa Excite by General Electric (Milwaukee, WI). The MEG data are collected using the 275-channel VSM system described earlier. The voxel size is 7 mm per side and $N_g \cong 7000$. For PFA denoising we use $N_s = 2$ and $N_x = 5$ and for ICA demixing we use AR-MOG with $K_g = 0$ and $K_Q = 3$.

Figure 5 shows the trial-averaged waveforms for 20 and 247 trials of data, where the trials are aligned according to the stimulus onset prior to averaging. Also shown in this figure are the two source estimates of the proposed method found for the case that there are only 20 trials. The M100 is visible in the source estimates of the proposed method with as few as 5 trials, whereas it takes 10 or 20 trials in order for the M100 to be evident in the trial-averaged data (not shown). Also, the proposed method indicates there is significant activity around 200 ms post-stimulus using as few as 20 trials, although this later activity cannot be seen in the trial-averaged data until there are much more than 50 trials.

Figure 6 shows the localization results of the PFA/Column Correlation method using the three standard orthogonal projections and the neurological convention. The voxels having the highest energy are highlighted, where the energy of each voxel is found by time-averaging over the post-stimulus period. Both coronal slices of the tomographic map produced by the proposed method show activation in the primary auditory cortex.

V. Discussion

The three methods, known to the authors, that are most similar to the proposed method are Bayesian methods by Knuth [30] and Roberts [43], and a convolutive ICA method by Took et al. [39]. These methods assume that the sources are statistically independent and they incorporate a physical forward model, as does the proposed method. The first of the methods assumes that the sources radiate spherically and the amplitude falls as the distance squared. This is unlike our model, where each source has an orientation (hence, radiation is not spherical) and the magnitude falls as the distance cubed. The second method imposes a non-negativity constraint on the ICA mixing matrix. The third method is designed for convolutive mixtures and it requires that the number of sensors is equal to or larger than the number of potential source locations. This is unlike the standard model for EEG and MEG data, which consists of an instantaneous mixture of the sources and where the number of putative source locations far exceeds the number of sensors.

Good denoising performance does not necessarily equate with good localization performance. For example, the mean PCA/Beamformer performance shown in Figures 2 and 4 is always worse than for the Raw Data/Beamformer even though PCA successfully removes noise (as measured by output SNR, the results of which are not shown here). For our data, PFA denoising outperforms PCA denoising when both use the same localization method (beamformer) and column correlation outperforms the beamformer when both use the same denoising method (PFA). The peaks in the tomographic map of the proposed method are not as sharp as the beamformer methods since the correlation of two columns of $\mathbf{F}_s$ decays slowly as the distance between the two associated voxels increases (especially if the orientations associated with the two voxels are approximately collinear). ICA denoising consistently suffers from over-fitting, which is not surprising since it must learn $N_b^2=75625$ parameters. Hence, the two methods that use ICA denoising, ICA/sLORETA and ICA/Dipole Fit, perform poorly. The proposed method, on the other hand, must learn approximately $N_b * N_g + N_s^2=554$ parameters. The ICA/sLORETA, ICA/Dipole Fit, and all 3 beamformer methods require the inversion of a $(N_b \times N_b)$ matrix ($N_b = 275$ for our data), whereas the proposed method requires the inversion of a $(N_s \times N_s)$ matrix ($N_s = 2$ for our data). Finally, the magnitude of the normalized correlation between the localization and separation metrics is 0.46, which indicates that the localization performance of the proposed method is not overly sensitive to the quality of the source separation.

Acknowledgments

This work was partially supported by NIH grant F32-NS052048 and NIH Grant R01-DC004855.

Biographies

Kenneth E. Hild II (M'91, SM'05) received a Ph.D. in Electrical Engineering from the University of Florida, Gainesville, where he studied information theoretic learning and blind source separation in the Computational NeuroEngineering Laboratory. Later, he studied biomedical informatics at Stanford University, Palo Alto. From 1995 to 1999, Dr. Hild worked in the Advanced Concepts group at Seagate Technologies. From 2000 to 2003, he taught several graduate-level classes on adaptive filter theory and stochastic processes at the University of Florida. From 2003 to 2007, he was at the Biomagnetic Imaging Laboratory at UCSF, San Francisco, where he applied variational Bayesian techniques for biomedical signal processing of encephalographic and cardiographic data. He is currently employed at the Oregon Health & Science University, Portland, where he is performing research on human and computer vision. His interests include full Bayesian models for competing versions of evolution, reinforcement learning using a hierarchy of innate concepts, and general-purpose machine learning algorithms.

Srikantan S. Nagarajan (S'90, M'91, SM'07) began his research career at the Applied Neural Control Laboratory of Case Western Reserve University (CWRU), where he obtained both an M.S. and a Ph.D. in Biomedical Engineering. After graduate school, he did a postdoctoral fellowship at the Keck Center for Integrative Neuroscience at the University of California, San Francisco, under the mentorship of Drs. Michael Merzenich and Christoph Schreiner. Subsequently, he was a tenure-track faculty member in the Department of Bioengineering at the University of Utah. Currently, he is the Director of the Biomagnetic Imaging Laboratory, an Associate Professor in Residence in the Department of Radiology, and a member in the UCSF/UCB Joint Graduate Program in Bioengineering. His research interests are in the area of Neural Engineering where his goal is to better understand dynamics of brain networks involved in processing, and learning, of complex human behaviors such as speech, through the development of functional brain imaging technologies.

References

- [1]. Jutten C, Herault J. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing* July;1991 24(1):1–10.
- [2]. Cardoso JF. Blind signal separation: Statistical principles. *Proc. IEEE* Oct;1998 86(10):2009–2025.

- [3]. Cardoso, JF. Super-symmetric decomposition of the fourth-order cumulant tensor: Blind identification of more sources than sensors; Intl. Conf. on Acoustics, Speech, and Signal Processing; Toronto, Canada. May 1991; p. 3109-3112.
- [4]. Yu Y, Petropulu AP. PARAFAC-Based Blind Estimation Of Possibly Underdetermined Convolutional MIMO Systems. *IEEE Trans. on Signal Proc* Jan;2008 56(1):111–124.
- [5]. Comon P, Rajih M. Blind identification of under-determined mixtures based on the characteristic function. *Signal Processing* Sept;2006 86(9):2271–2281.
- [6]. Atashbar, M.; Kahaei, MH. Information and Communication Technologies. Vol. 1. Damascus, Syria: Apr. 2006 Multi-Speakers Direction-Of-Arrival Finding using ICA; p. 1212-1213.
- [7]. Raju K, Ristaniemi T, Karhunen J, Oja E. Jammer Suppression in DS-CDMA Arrays Using Independent Component Analysis. *IEEE Trans. on Wireless Commun* Jan;2006 5(1)
- [8]. Bell AJ, Sejnowski TJ. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation* Nov;1995 7(6):1129–1159. [PubMed: 7584893]
- [9]. Pham DT. Mutual information approach to blind separation of stationary sources. *IEEE Trans. on Information Theory* July;2002 48(7):1935–1946.
- [10]. Kwak, N.; Choi, CH.; Choi, JY. Feature Extraction Using ICA. *Lecture Notes in Computer Science; Intl. Conf. on Artificial Neural Networks; Vienna, Austria. Aug. 2001; p. 568-573.* Springer, Berlin
- [11]. Vigario, R.; Jousmaki, V.; Hamalainen, M.; Hari, R.; Oja, E. Independent component analysis for identification of artifacts in magnetoencephalographic recordings; *Advances in Neural Information Proc. Systems; MIT Press, Cambridge, MA. Nov. 1994; p. 229-235.*
- [12]. James CJ, Gibson OJ. Temporally constrained ICA: An application to artifact rejection in electromagnetic brain signal analysis. *IEEE Trans. on Biomedical Engineering* Sept;2003 50:1108–1116.
- [13]. Chen, L.; Wang, C.; Shih, IM.; Wang, TL.; Zhang, Z.; Wang, Y.; Clarke, R.; Hoffman, E.; Xuan, J. Biomarker Identification by Knowledge-Driven Multi-Level ICA and Motif Analysis; *Intl. Conf. on Machine Learning and Applications; Cincinnati, Ohio. Dec. 2007; p. 560-566.*
- [14]. Dawy Z, Sarkis M, Hagenauer J, Mueller JC. Fine-Scale Genetic Mapping Using Independent Component Analysis. *IEEE/ACM Trans. on Com. Biology and Bioinformatics* July-Sept;2008 5:448–460.
- [15]. Naik, GR.; Kumar, DK.; Weghorn, H. ICA based identification of sources in sEMG; *Intl. Conf. on Intelligent Sensors, Sensor Networks and Information; Melbourne, Australia. Dec. 2007; p. 619-624.*
- [16]. Sarvas J. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine and Biology* Jan;1987 32(1):11–22. [PubMed: 3823129]
- [17]. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* Mar;2004 134:9–21. [PubMed: 15102499]
- [18]. Dalal SS, Zumer JM, Agrawal V, Hild KE, Sekihara K, Nagarajan SS. NUTMEG: A neuromagnetic source reconstruction toolbox. *Neurology and Clinical Neurophysiology* Nov;2004 52:1–4.
- [19]. Sekihara, K.; Nagarajan, S. *Adaptive Spatial Filters for Electromagnetic Brain Imaging.* Springer; NY: Jun. 2008
- [20]. Sekihara K, Nagarajan S, Poeppel D, Marantz A. Asymptotic SNR of scalar and vector minimum-variance beamformers for neuromagnetic source reconstruction. *IEEE Trans. on Biomedical Engineering* Oct;2004 51(10):1726–1734.
- [21]. Mosher JC, Lewis PS, Leahy RM. Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Trans. on Biomedical Engineering* June;1992 39(6):541–557.
- [22]. Ossadtchi A, Baillet S, Mosher JC, Thyerlei D, Sutherling W, Leahy RM. Automated interictal spike detection and source localization in magnetoencephalography using independent components analysis and spatio-temporal clustering. *Clinical Neurophysiology* Mar;2004 115(3):508–522. [PubMed: 15036046]
- [23]. Tang AC, Pearlmutter BA, Malaszenko NA, Phung DB, Reeb BC. Independent components of magnetoencephalography: Localization. *Neural Computation* Aug;2002 14:1827–1858. [PubMed: 12180404]

- [24]. Leahy RM, Mosher JC, Spencer ME, Huang MX, Lewine JD. A study of dipole localization accuracy for MEG and EEG using a human skull phantom. *Electroencephalography and Clinical Neurophysiology* April;1998 107(2):159–173. [PubMed: 9751287]
- [25]. Kobayashi K, Akiyama T, Nakahori T, Yoshinaga H, Gotman J. Systematic source estimation of spikes by a combination of independent component analysis and RAP-MUSIC: I: Principles and simulation study. *Clinical Neurophysiology* May;2002 113:713–724. [PubMed: 11976051]
- [26]. Grosse-Wentrup, M.; Gramann, K.; Wascher, E.; Buss, M. EEG source localization for brain-computer-interfaces; *IEEE Engineering in Medicine and Biology Society Conference on Neural Engineering*; Arlington, VA. Mar. 2005; p. 128-131.
- [27]. Pascual-Marqui RD. Review of Methods for Solving the EEG Inverse Problem. *Intl. Journal of Bioelectromagnetism* 1999;1:75–86.
- [28]. Niedermeyer, E.; Lopes da Silva, F. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Vol. 5th ed.. Lippincott Williams & Wilkins; Philadelphia, PA: 2004. p. 280
- [29]. Knuth, K. *Maximum Entropy and Bayesian Methods*. Munich, Germany: 1998. Convergent Bayesian formulations of blind source separation and electromagnetic source estimation; p. 217-226.
- [30]. Knuth, KH. Bayesian source separation and localization; *Proc. of SPIE, Bayesian Inference for Inverse Problems*; San Diego, CA, SPIE Press, Bellingham, WA. July 1998; p. 147-158.
- [31]. Hild, KE., II; Nagarajan, S. Source localization of EEG/MEG data by correlating columns of ICA solution with lead field matrix; *Intl. Conf. on Acoustics, Speech, and Signal Processing*; Honolulu, Hawaii. April 2007; p. 1177-1180.
- [32]. Cichocki, A.; Amari, S. *Adaptive blind signal and image processing*. John Wiley & Sons; Chichester, England: 2003. p. 305-309.
- [33]. Mosher JC, Leahy RM, Lewis PS. EEG and MEG: Forward Solutions for Inverse Methods. *IEEE Trans. on Biomedical Engr* Mar;1999 46(3):245–259.
- [34]. Grosse-Wentrup M, Buss M. Multiclass Common Spatial Patterns and Information Theoretic Feature Extraction. *IEEE Trans. on Biomedical Engineering* Aug;2008 55(8):1991–2000.
- [35]. Nagarajan, S.; Attias, HT.; Sekihara, K.; Hild, KE, II. Independent Component Analysis and Signal Separation. *Charleston, SC: Mar. 2006 Partitioned factor analysis for interference suppression and source extraction*; p. 189-197.
- [36]. Nagarajan, Srikantan S.; Attias, Hagai T.; Hild, Kenneth E., II; Sekihara, Kensuke. A probabilistic algorithm for robust interference suppression in bioelectromagnetic sensor data. *Statistics in Medicine* Sept 20;2007 26(21):3886–3910. [PubMed: 17546712]
- [37]. Adjamiana P, Barnes GR, Hillebrand A, Holliday IE, Singha KD, Furlong PL, Harringtonb E, Barclayb CW, Route PJG. Co-registration of magnetoencephalography with magnetic resonance imaging using bite-bar-based fiducials and surface-matching. *Clinical Neurophysiology* Mar;2004 115(3):691–698. [PubMed: 15036065]
- [38]. Hild KE II, Attias HT, Nagarajan S. An Expectation-Maximization method for spatio-temporal blind source separation using an AR-MOG source model. *IEEE Trans. on Neural Networks* Mar; 2008 19:508–519.
- [39]. Took CC, Sanei S, Rickard S, Chambers J, Dunne S. Fractional Delay Estimation for Blind Source Separation and Localization of Temporomandibular Joint Sounds. *IEEE Trans. on Biomedical Engr* Mar;2008 55(3):949–956.
- [40]. Knuth, K. *Maximum Entropy and Bayesian Methods*. Boise, ID: Aug. 1997 Difficulties applying recent blind source separation techniques to EEG and MEG; p. 209-222.
- [41]. Weinstein, Ehud; Feder, Meir; Oppenheim, Alan. Multi-channel signal separation by decorrelation. *IEEE Trans. on Speech and Audio Proc* Oct;1993 1(4):405–413.
- [42]. Parra, Lucas; Spence, Clay. Convolutional blind separation of non-stationary sources. *IEEE Trans. on Speech and Audio Proc* May;2000 8(3):320–327.
- [43]. Roberts, S.; Choudrey, R. *Deterministic and Statistical Methods in Machine Learning*. Vol. 3635. Springer; Berlin: 2005. Bayesian independent component analysis with prior constraints: An application in biosignal analysis; p. 159-179.

Fig. 1.

Contour plots of the tomographic maps for a single Monte Carlo trial. The •'s and ×'s correspond to the true neural source and interference source locations, respectively. Results are shown for Rawdata/Beamformer (upper left), PFA/Beamformer (upper right), PCA/Beamformer (lower left), and PFA/Column Correlation (lower right).

Fig. 2.
Mean localization error of the 1st experiment as a function of SIR.

Fig. 3.

Contour plots of the tomographic maps for a single Monte Carlo trial. The •'s and ×'s correspond to the true neural source and interference source locations, respectively. Results are shown for Rawdata/Beamformer (upper left), PFA/Beamformer (upper right), PCA/Beamformer (lower left), and PFA/Column Correlation (lower right).

Fig. 4.
Mean localization error of the 2nd experiment as a function of SIR.

Fig. 5. Trial-averaged data computed using the first 20 trials (upper) and using all 247 trials (middle). The two estimated sources, for the proposed method, computed using 20 trials of data (lower).

Fig. 6. Source localization result for (real) auditory stimulus-evoked data using the 1st component of PFA/Correlate Column. Axial, sagittal, and coronal (left side) views are shown.