

A NON-UNIFORMLY SAMPLED MARKOV RANDOM FIELD MODEL FOR MAP RECONSTRUCTION OF MAGNETOENCEPHALOGRAPH IMAGES *

Alan H. Gardiner† and Brian D. Jeffs‡

† Lockheed Martin Federal Systems

‡ Department of Electrical and Computer Engineering, 459 CB, Brigham Young University, Provo, UT 84602

ABSTRACT

In this paper the maximum *a posteriori* (MAP) image reconstruction of magnetoencephalograms (MEG) is investigated. A mathematical framework for vector Markov random field models (MRF) suitable for MEG modeling of brain neuron current dipole activity is developed. A new method for simulating an MRF over a non-uniformly spaced sample grid while approximating an arbitrary desired covariance structure at these samples is also presented. Simulation results validate the effectiveness of this random sampled field model, and clinical MEG evoked response data is processed to demonstrate algorithm performance.

1. INTRODUCTION

In this paper the maximum *a posteriori* (MAP) reconstruction of magnetoencephalograms (MEG) is investigated. A simple linear relationship between source current dipoles arising from neuron electrical activity and the resulting extracranial magnetic fields is used so that the solution may be cast as a classical inverse imaging problem. A mathematical framework for Markov random field models (MRF) suitable for MEG will be developed.

A new method for simulating an MRF over an arbitrary (non-uniform) sample grid is also presented. In this new technique the covariance matrix may be specified by the analyst to be consistent with expected spatial correlations in neural activity. The neighborhood weighting coefficients for the non-uniformly sampled Gaussian Markov random field (GMRF) are then computed to produce the desired covariance at sample locations. Using these techniques, MAP reconstructions onto a hemispheric brain model are performed using simulated and real MEG data.

Neuromagnetic imaging (NMI), the process of creating a source image consistent with the magnetic field observations, was first performed on a single plane with constrained source orientations by Singh *et al.* [1]. Currently, one of the leading areas of research in NMI is in the Bayesian context, with MRF's used to represent any prior knowledge about source the distribution in the form of an image prior probability density function (pdf). In the NMI framework,

*THE AUTHORS EXPRESS APPRECIATION TO DR. JOHN MOSHER OF LOS ALAMOS NATIONAL LABORATORY AND THE VETERAN'S ADMINISTRATION HOSPITAL IN ALBUQUERQUE, NM FOR PROVIDING THE MEG CLINICAL DATA USED IN THIS PAPER.

Phillips *et al* have developed a composite, dual MRF model using a binary indicator process and independent Gaussian random variables to represent whether a source is on or off and source amplitude respectively [2]. Their model assumes a prior knowledge of source orientation, so a scalar MRF may be used.

In our proposed method, a single consistent vector field MRF source model is introduced. The associated potential function is designed to both encourage solution sparseness and to permit MRF modeling on the non-uniformly sampled hemispherical surface used to represent the brain cortex. Because this MRF models vector fields, dipole orientations are included in the model and are not required to be known *a priori*.

2. SENSOR SYSTEM RESPONSE MODEL

Neuron electrical activity in the brain will be modeled as a distribution of discrete current dipoles, \mathbf{Q}_s , each with an unknown magnitude and 3-D orientation. The goal of MEG imaging is to estimate the values of these dipole parameters on a sample grid throughout the tissue of interest. Neuron current dipole activity is observed indirectly using an array of sensitive magnetometers placed around the skull. These sensors employ Superconducting QUantum Interference Device (SQUID) circuitry coupled with a pickup coil wound in a gradiometer configuration, so as to be sensitive to locally induced magnetic fields only.

We will use a linear observation model for data from the array of MEG sensors. We define the 3×1 element transfer vector, $\Gamma_{i,s}$, to satisfy the static Biot-Savart law for single sensor response to a single active dipole,

$$b_i = \Gamma_{i,s} \cdot \mathbf{Q}_s = \frac{\mu_0 (\mathbf{r}_i - \mathbf{r}_s) \times \hat{v}(i)}{4\pi |\mathbf{r}_i - \mathbf{r}_s|^3} \cdot \mathbf{Q}_s \quad (1)$$

where \cdot and \times represent vector inner and outer products respectively, b_i is the i^{th} MEG sensor output, \mathbf{r}_i is its 3-D position vector, $\hat{v}(i)$ is the unit direction vector normal to the plane containing the sensor's gradiometer coil, \mathbf{Q}_s is the 3-D current dipole vector at the s^{th} pixel site, and \mathbf{r}_s is the position vector for this dipole. The entire array observation may be represented in matrix-vector product form

$$\begin{aligned} \mathbf{b} &= \mathbf{G}\mathbf{q} + \boldsymbol{\eta}, \\ \mathbf{b} &= [b_1, \dots, b_M]^T, \end{aligned} \quad (2)$$

$$\mathbf{G} = \begin{bmatrix} \Gamma_{1,1}^T & \cdots & \Gamma_{1,S}^T \\ \vdots & & \vdots \\ \Gamma_{M,1}^T & \cdots & \Gamma_{M,S}^T \end{bmatrix},$$

$$\mathbf{q} = [\mathbf{Q}_1, \dots, \mathbf{Q}_S]^T,$$

where η represents sensor error (noise), and M and S are the total number of sensor and pixel sites respectively. Figure 1 shows the actual sensor geometry for the Neuromag model 122 SQUID detector head array that we will use. There are 122 conformal gradiometers, with two mutually orthogonal loop pairs at each of the 61 sites. The brain and skull are modeled as a conducting sphere encased in an insulating shell. No radial components of the dipole mo-

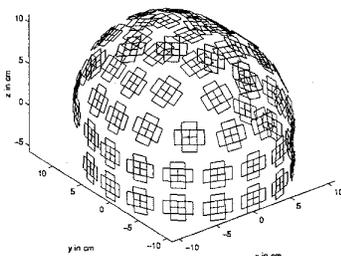


Figure 1. Plot showing gradiometer locations and configurations for the Neuromag 122 system.

ments can be detected under the conducting sphere model, so we assume these terms are zero. This assumption is incorporated into our model by restricting current dipoles to be oriented tangentially to a hemispherical surface which corresponds to the brain cortex, and lies about 3 cm below the sensor array shell.

The dipole sample grid (i.e. set of all possible current dipole sites) must have non-uniform spacing due to the spherical surface geometry, and therefore our non-uniformly sampled MRF model is called for. Due to the tangential constraint, and with the dipole hemisphere centered at the origin, dipoles may be completely described by magnitude and azimuth, given their latitude (elevation angle) and longitude on the hemisphere to locate the sample. Let $\bar{\mathbf{q}} = [\bar{q}_1, \dots, \bar{q}_S]^T$ be the vector of dipole magnitudes, with $\bar{q}_s = |\mathbf{Q}_s|$, and let $\Theta = [\theta_1, \dots, \theta_S]^T$ be the corresponding vector of tangentially constrained dipole azimuths. An azimuth of zero is defined as the vector orientation tangent to the hemisphere that would have the largest positive projection onto the positive z axis, i.e. zero azimuth vectors point toward the z axis. The s^{th} dipole vector element of \mathbf{q} , as a function of \bar{q}_s and θ_s can be expressed in Cartesian coordinates as

$$\mathbf{Q}_s = \begin{bmatrix} -\sin(\lambda_s) & -\sin(\phi_s) \cos(\lambda_s) \\ \cos(\lambda_s) & -\sin(\phi_s) \sin(\lambda_s) \\ 0 & \cos(\phi_s) \end{bmatrix} \begin{bmatrix} \bar{q}_s \sin \theta_s \\ \bar{q}_s \cos \theta_s \end{bmatrix} \quad (3)$$

where $\mathbf{r}_s = \lambda_s$ is the elevation angle (angle relative to the $x - y$ plane) of the s^{th} sample site, and ϕ_s is the longitude (angle relative to the $x = z$ plane). Equation (3) will be used in Section 4. as a constraint in our non-uniformly sampled vector field MRF dipole model.

3. NON-UNIFORMLY SAMPLED GAUSS-MARKOV RANDOM FIELDS

MEG imaging is known to be a notoriously ill posed inverse problem, with a very large ambiguity in the solution space. The proposed algorithm is based on the Bayesian MAP image restoration method because of the natural means it provides for regularizing the inverse solution. The MAP estimate of \mathbf{q} is given by

$$\hat{\mathbf{q}} = \arg \min_{\mathbf{q}} f_{Q|B}(\mathbf{q}|\mathbf{b}) = \arg \min_{\mathbf{q}} f_{B|Q}(\mathbf{b}|\mathbf{q}) f_Q(\mathbf{q}) \quad (4)$$

where $f(\cdot)$ represents a probability density function (pdf). The image prior pdf, $f_Q(\mathbf{q})$, must be chosen to model statistical characteristics of the desired true image, so as to yield a unique solution. We will introduce a new non-uniformly sampled GMRF model for $f_Q(\mathbf{q})$ which significantly improves reconstruction results.

In the absence of detailed brain cortical structure information, and since brain tissue is homogeneous at the scale used for the dipole grid, we are obliged to adopt a homogeneous model for the covariance structure of the current dipole distribution. We will assume that the underlying physical current distribution in brain tissue is a stationary, zero mean, continuous Markov random field [3] with an isotropic covariance structure.

This underlying stationary continuous MRF is sampled at arbitrary points to provide elements of a discrete MRF for our reconstruction model. The correlations between these samples are set to match those of the underlying continuous field, but the resulting discrete MRF will in general be neither stationary (with respect to the ordering of sample indices) nor isotropic, since samples at a uniform spacing, r_0 , will likely not be included in the discrete model. Note again that uniform sampling is impossible (and probably not desirable) on a hemispherical surface.

A stationary discrete random field is easily synthesized using Markov random field models on a uniform lattice, but we must be able to generate a MRF on a random grid overlying the isotropic, homogeneous, continuous covariance model. Let $R_q(\psi, \tau)$ be the covariance function between two continuous space points, ψ and τ , on the dipole hemisphere. By isotropic, we mean that $R_q(\psi, \tau)$ is only a function of the Euclidean distance between ψ and τ . Now if s and t are points on the non-uniformly sampled lattice, \mathcal{S} , then we require that the covariance between these samples in the discrete MRF, \mathbf{q} , be given by $E\{\mathbf{Q}_s \cdot \mathbf{Q}_t\} = R_q(\psi = \mathbf{r}_s, \tau_t = \mathbf{r}_t)$.

We will first develop a method for approximating a desired covariance model on a non-uniformly sampled grid for scalar MRF's. This will be extended to vector field models in the following section. The joint log likelihood function of a scalar Gauss-Markov random field, \mathbf{x} can be expressed as [4, 5]

$$\log f_X(\mathbf{x}) \propto - \left(\sum_{s \in \mathcal{S}} a_s x_s^2 + \sum_{\{s,t\} \in \mathcal{C}} b_{st} |x_s - x_t|^2 \right) \quad (5)$$

Neighborhood influence parameters (potential weights), b_{st} and a_s in (5) must be chosen to model the desired covariance of the continuous spatial distribution. In uniformly

sampled MRF's, this is easily accomplished given a desired covariance matrix for \mathbf{x} , \mathbf{R}_x . However, no such method is available in the literature for non-uniformly sampled fields, and it is impossible to sample a hemisphere with a uniform lattice. Other authors go about choosing a Markov model by specifying fixed a_s and b_{st} values for their homogeneous uniformly sampled fields regardless of the site s [4, 5, 6]. In the case where the sampling is somewhat less than uniform, this model does not produce homogeneous images because the potential function in (5) provides no means of reducing the influence of distant neighbors or increasing the influence of near neighbors. Perhaps a more important issue that arises under non-uniform sampling is the choice of sites to include in the neighborhood of each pixel. Heuristically, those members of the neighborhood δ_s which are very near to s should exact a greater penalty for dissimilarity than those neighbors which are more distant. The trouble with this intuitive argument is that the exact mathematical relationship between the neighborhood weighting coefficients and the sample distance is not readily apparent.

Equation (5) can also be represented as a multivariate jointly normal density with covariance matrix $\mathbf{R}_x = \mathbf{B}^{-1}$, where interaction matrix \mathbf{B} has elements, $\{B_{st}\}$, which relate to our GMRF influence parameters as follows: $a_s = \sum_{t \in \delta_s} B_{st}$ and $b_{st} = -B_{st}$ [4]. This appears to give us a direct method to calculate neighborhood influence parameters from a specified covariance matrix. In a 1-D MRF, an exponential covariance function yields a tri-diagonal \mathbf{B} , where most $B_{st} = 0$. Unfortunately, with non-uniform sampling, even the simplest of 2-D continuous covariance functions, $R_x(\psi, \tau)$, require nearly all the B_{st} to be non-zero [7]. An MRF model is only computationally useful if the neighborhood system is small, i.e. if each pixel has only a few neighbors. With most $B_{st} \neq 0$, every pixel is a neighbor of every other pixel, and the MRF model is useless for reconstruction algorithms. We will present a method for approximating \mathbf{B} for a specified $R_x(\psi, \tau)$ such that the neighborhood system remains small, and the model matches the desired correlation structure much more closely than if conventional uniform sample grid methods were used.

The goal is to find an approximation, $\hat{\mathbf{B}}$, for \mathbf{B} such that: a) most of the $\hat{B}_{st} = 0$, i.e. the neighborhood structure is small, b) $\|\hat{\mathbf{B}}^{-1} - \mathbf{R}_x\|_F \leq \epsilon$, i.e. the approximation error is bounded, c) the element-wise variances of \mathbf{x} match the desired values, and d) $\hat{\mathbf{B}}$ is positive definite, inevitable. Positive definiteness is required for the corresponding probability distribution to be well defined. We have shown that the following approximation meets these criteria [7]

$$\begin{aligned} \hat{\mathbf{B}} &= \mathbf{D}^{\frac{1}{2}} \mathbf{A} \mathbf{D}^{\frac{1}{2}}, \\ \mathbf{A} &= \mathbf{M} \odot \mathbf{R}_x^{-1} + \mathbf{E}, \\ \mathbf{D} &= \text{diag}\{\mathbf{A}\} \text{diag}\{\mathbf{R}_x\} \end{aligned} \quad (6)$$

where \odot indicates the Schur element-by-element matrix product, $\text{diag}\{\}$ is a diagonal matrix formed from the diagonal of the argument. \mathbf{M} is a selection matrix which controls the size of the neighborhood for each pixel, and has elements m_{st} given by

$$m_{st} = \begin{cases} 1 & t \in \delta_s \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

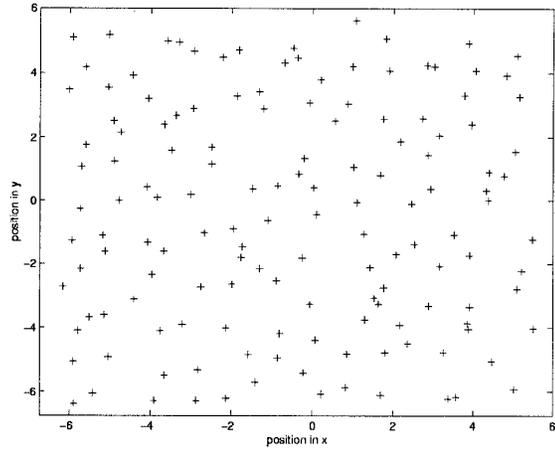


Figure 2. Pseudo-randomly sampled field used in the experiment of Figure 3.

δ_s is the set of all sites which are neighbors of pixel s . A simple threshold test is used to form this set. For a given pixel s , t is a neighbor of s if $|B_{st}| \geq \tau$. τ is manually selected to be just large enough to achieve the desired neighborhood size, and thus the order of the approximation. \mathbf{E} is a diagonal matrix with elements $e_{ss} = \sum_{t \notin \delta_s} B_{st}$. The effect of \mathbf{E} is to insure $\hat{\mathbf{B}}$ will be positive definite and to assist in keeping the pixel marginal variances at the desired level.

It may be noted that forming $\hat{\mathbf{B}}$ requires computing the inverse of a potential very large matrix, \mathbf{R}_x . In the MEG imaging problem addressed here, the number of voxels in the image space is relatively small (on the order of 100's of voxels), so this issue is not a problem. On the other hand, even when the number of pixels is very large, $\hat{\mathbf{B}}$ need be evaluated only once for a given imaging geometry, and all subsequent reconstructions can use the same pre-computed value. Further simplification is possible by noting that typically $R_x(\psi, \tau)$ drops off rapidly with distance, so neighborhood influence parameters, b_{st} for distant pixels will undoubtedly be set to zero. The image field, and correspondingly \mathbf{R}_x and $\hat{\mathbf{B}}$ can therefore be partitioned into smaller, overlapping regions such that the inverses are easy to compute.

The effectiveness of the above approximation was demonstrated on a non-uniformly sampled Gauss-Markov random field. A simple 2-D isotropic exponential covariance was used for the underlying continuous field,

$$R_x(\psi, \tau) = \rho^{||\psi - \tau||_2}. \quad (8)$$

The corresponding desired discrete interaction matrix, \mathbf{B} , is given by

$$\begin{aligned} \{\mathbf{B}^{-1}\}_{st} &= \{\mathbf{R}_x\}_{st} = R_x(\psi = \mathbf{r}_s, \tau = \mathbf{r}_t) \\ &= \rho^{||\mathbf{r}_s - \mathbf{r}_t||_2}. \end{aligned} \quad (9)$$

Figure 2 shows positions of 144 random samples sites which overlay the exponential correlation continuous MRF. Using these sample locations, the approximation of equation (6)

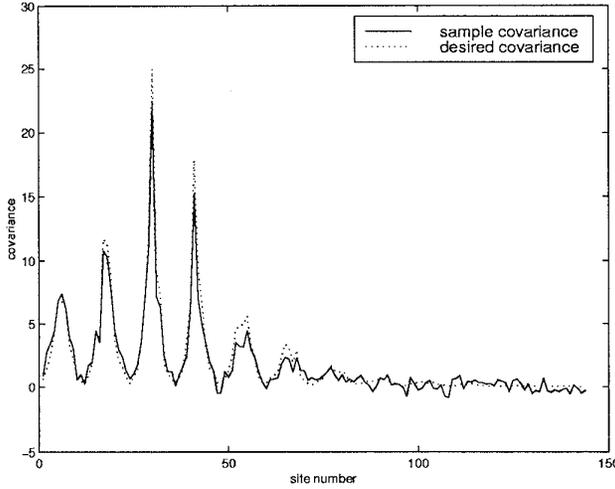


Figure 3. Sample covariance vs. desired covariance for an interior pixel in the field with high sampling randomness.

was applied to compute the GMRF neighborhood weights, with $\rho = 0.5$, and τ set to give approximately four neighbors for each pixel site. Chen's algorithm [8] was used to generate 2000 realizations of the GMRF random field so that sample covariances could be computed between pixel sites and compared with the desired values given by \mathbf{R}_x . Figure 3 compares covariance values (relative to an interior pixel at $s = 30$) for a synthesized non-uniform GMRF with the corresponding desired values. Note that the synthesized data has sample cross covariance values that track the ideal desired values remarkably closely, with a model neighborhood the same size as a uniformly sampled first order neighborhood. Experiments (not shown) using the best match uniform weighting (i.e. as in uniform rectilinear sampling) produced dramatically higher error.

4. EXTENSION TO VECTOR FIELD MODELS FOR MEG RECONSTRUCTION

The vector GMRF model is an obvious extension of scalar equation (5). Note that the second term in this equation is a penalty (in the exponent of $f_x(\mathbf{x})$) for dissimilar intensity values between pixels x_s and x_t . The penalty is proportional to the square of the difference. For a vector field, we wish to penalize (in a square law sense) differences in the vectors at sites s and t . This penalty should encourage neighboring dipoles to have similar magnitudes and orientations. We propose the following model, which simply replaces scalar arithmetic in equation (5) with the corresponding 3-D vector operations,

$$\log f_q(\mathbf{q}) \propto \sum_{s,t \in \mathcal{C}} b_{st} (\mathbf{Q}_s - \mathbf{Q}_t)^T (\mathbf{Q}_s - \mathbf{Q}_t) + \sum_{s \in \mathcal{S}} a_s \mathbf{Q}_s^T \mathbf{Q}_s. \quad (10)$$

On the cortex hemisphere sample surface, it is not possible in general for two sample sites to have exactly the same arbitrary dipole orientation, because the two corresponding constraint tangent planes are not necessarily parallel. However, since only local neighbors are considered in

the approximation model, the hemisphere surface curvature between neighbor pixels is small. The tangent constrained dipoles can have approximately equal vectors. This preference in the model for neighboring dipole vectors to be as similar as possible within the orientation constraints agrees with our notions of what a correlated vector field should be.

Assuming white Gaussian noise, adopting the vector GMRF model of equation (10) for $f_Q(\mathbf{q})$, and confining dipoles to lie tangential to the cortex surface hemisphere, leads to the MAP reconstruction solution

$$\hat{\mathbf{q}} = \arg \min_{\hat{\mathbf{q}}, \Theta} (\mathbf{b} - \mathbf{G}\hat{\mathbf{q}})^T (\mathbf{b} - \mathbf{G}\hat{\mathbf{q}}) + \lambda \sum_{s,t \in \mathcal{C}} b_{st} (\mathbf{Q}_s - \mathbf{Q}_t)^T (\mathbf{Q}_s - \mathbf{Q}_t) + \lambda \sum_{s \in \mathcal{S}} a_s \mathbf{Q}_s^T \mathbf{Q}_s \quad (11)$$

where \mathbf{Q}_s and \mathbf{Q}_t are evaluated using equation (3) as a constraint, and λ is the regularization weight which controls the relative influence on the solution of the image prior $f_Q(\mathbf{q})$. Though in theory λ could be calculated exactly given $\hat{\mathbf{B}}$, this is difficult, and in practice λ is adjusted manually for best restoration results. \mathcal{C} is the set of all pixel cliques in the MRF neighborhood system, and \mathcal{S} is the set of all sample points on the dipole hemisphere lattice. This formulation does an excellent job of producing realistic solutions such that neighboring dipoles have similar orientations (azimuths). Equation (11) can be solved using a stochastic search algorithm (e.g. Metropolis), or a simple gradient descent technique.

5. RESULTS WITH CLINICAL DATA

Figure 5 shows results of image reconstruction using equation (11) on clinical data collected using the Neuromag 122 array. This data represents the somatosensory evoked response from piezoelectric stimulation of the right ring finger. Data is averaged over 109 trials to increase SNR, which even after averaging is less than 10 dB. Figure 4 presents an aggregate plot of the averaged data from all 122 sensors overlaid on a single axis. The sampling interval is 0.8 ms, and the spikes at about sample 125 are induced noise from the shock stimulus. The evoked response occurs at about sample 180. Note the poor SNR even in this averaged data, which makes the results of Figure 5(a,b) the more impressive.

Figure 5 shows two reconstructions of the data. The reconstruction was formed using a single time sample (for each sensor) 36 ms after the stimulus. This time corresponded to the largest evoked signal response. 5(c,d) depict a reconstruction with a smaller weighting (λ) on the image prior term. The active region of the brain for this experiment should be the left somatosensory cortex, located in the region above the left ear canal. Figure 5(a,b), shows a strong preference of our model to select this region. On the other hand, less dependence on the irregular lattice GMRF image prior creates an image which is not as definitive, as shown in Figure 5(c,d). Both solutions are equally consistent with the observed data, \mathbf{b} , which illustrates the need for a strong regularizing term in the reconstruction.

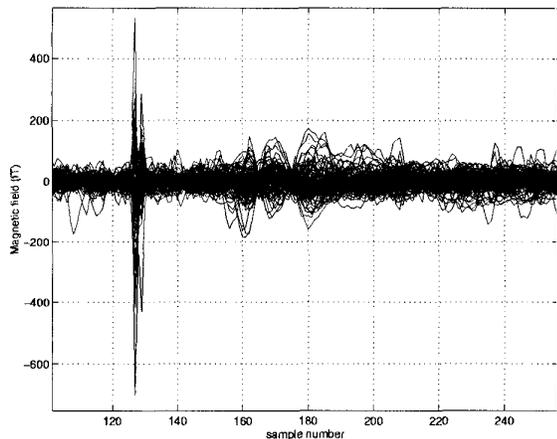


Figure 4. Plot of MEG data evoked by right ring finger stimulus.

6. CONCLUSIONS

This paper has demonstrated the effectiveness of a vector form Gauss-Markov random field model for Maximum *a posteriori* reconstruction, from external magnetic field measurements, of brain current dipole distributions. A technique was presented for achieving (approximately) an arbitrary desired covariance structure for a non-uniformly sampled Markov random field while keeping the neighborhood structure small. The experiments with real MEG data are encouraging. Solutions have reduced random noise, reduced canceling dipole content (i.e. dipole configuration components that are invisible at the sensors), and better resolution of isolated dipoles than an unregularized image reconstruction. These results suggest that further development of the technique is warranted.

REFERENCES

- [1] M. Singh, D. Doria, V.W. Henderson, G.C. Huth, and J. Beatty, "Reconstruction of images from neuromagnetic fields", *IEEE Transactions on Nuclear Science*, vol. NS-31, pp. 585-589, 1984.
- [2] J. W. Phillips, R. M. Leahy, and J. C. Mosher, "MEG-based imaging of focal neuronal current sources", *IEEE Transactions on Medical Imaging*, vol. 16, no. 3, pp. 338-48, June 1997.
- [3] J. W. Woods, "Two-dimensional discrete (markovian) fields", *IEEE Transactions on Information Theory*, pp. 232-240, Mar. 1972.
- [4] Wai Ho Pun, *Parametric Model-Adaptive Image Restoration*, PhD thesis, Brigham Young University, 1995.
- [5] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation", *IEEE Trans. on Image Processing*, vol. 2, no. 3, pp. 296-310, July 1993.
- [6] J. M. F. Moura and N. Balram, "Recursive structure of noncausal Gauss-Markov random fields", *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 334-54, Mar. 1992.

- [7] Alan H. Gardiner, "Map reconstruction of magnetoencephalograms", Master's thesis, Brigham Young University, 1997.
- [8] C. C. Chen, *Markov random field models in image analysis*, PhD thesis, Michigan State University, East Lansing, 1988.

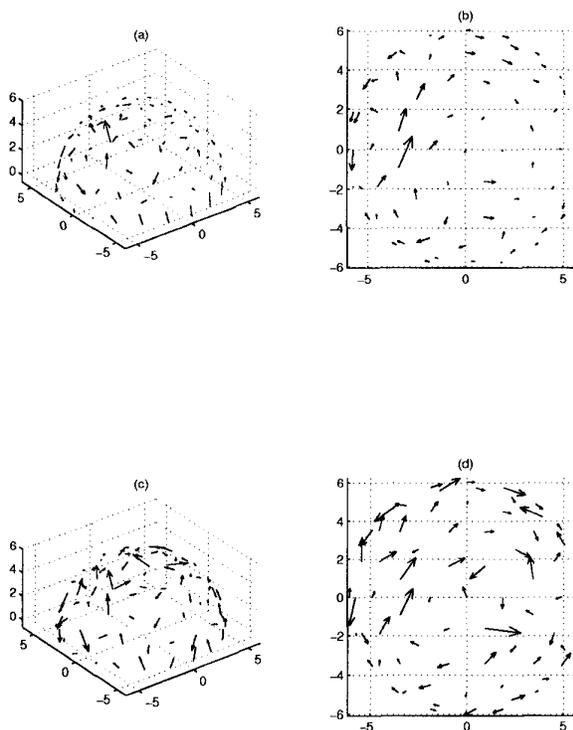


Figure 5. (a,b) Side and top views of vector GMRF reconstruction of clinical MEG data and (c,d) reconstruction with a lesser emphasis on the GMRF prior (i.e. λ smaller than in (a,b)). The top of (b) and (d) corresponds to the forward section of the brain.