Published in final edited form as:
Proc IEEE Int Symp Biomed Imaging. 2009 August 7; 2009: 149–152. doi:10.1109/ISBI.2009.5193005.

GENERALIZED SIDELOBE CANCELLER FOR MAGNETOENCEPHALOGRAPHY ARRAYS

John C. Mosher[1], **Matti S. Hämäläinen**[2], **Dimitrios Pantazis**[3], **Hua Brian Hui**[3], **Richard C. Burgess**[1], and **Richard M. Leahy**[3]

[1] Cleveland Clinic Neurological Institute, Epilepsy Center, Cleveland, OH USA

[2] Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA USA

[3] Signal & Image Processing Institute, University of Southern California, Los Angeles, CA USA

Abstract

In the last decade, large arrays of sensors for magnetoencephalography (MEG) (and electroencephalography (EEG)) have become more commonplace, allowing new opportunities for the application of beamforming techniques to the joint problems of signal estimation and noise reduction. We introduce a new approach to noise cancellation, the generalized sidelobe canceller (GSC), itself an alternative to the linearly constrained minimum variance (LCMV) algorithm. The GSC framework naturally fits within the other noise reduction techniques that employ real or virtual reference arrays. Using expository human subject data with strong environmental and biological artifacts, we demonstrate a straightforward sequence of steps for practical noise filtering, applicable to any large array sensor design.

Index Terms

Array signal processing; magnetoencephalography; spatial filters; adaptive arrays

1. INTRODUCTION

We review briefly the linear model, approximations, and statistical assumptions commonly used in MEG processing [1], [4], [6]. Our fundamental linear model is simply

$$\boldsymbol{d}=\boldsymbol{L}\boldsymbol{j}+\boldsymbol{n} \tag{1}$$

where $\boldsymbol{d}$ is the $m \times 1$ vector of measured data, measured at m external channels for 1 time sample, $\boldsymbol{L}$ is the $m \times p$ matrix representing the *lead-field model*, relating m external measurements to the amplitudes of p current dipoles throughout the source volume of interest, $\boldsymbol{j}$ is the $p \times 1$ vector of the signed amplitudes of the dipoles, and $\boldsymbol{n}$ is the $m \times 1$ vector which represents additive "noise" (anything not explicitly in the model $\boldsymbol{Lj}$).

For each sensor in the array, we calculate the lead field matrix $\boldsymbol{L}$ for a dense mesh of thousands of current dipoles within the brain volume (cf. [7] for a review and discussion on head model calculations). The number of sensors is about 100–300, and $\boldsymbol{L}$ therefore represents a system of highly underconstrained system of equations, $m \ll p$.

Because MEG and EEG are extracranial measurements made at a distance from the neural sources, fundamental quasistatic electromagnetics limit the amount of information that can be

recorded by these roughly helmet-like arrays of sensors. The density of the sensors is effectively redundant, such that $\boldsymbol{L}$ has a very high condition number, and we can in general create a surrogate overconstrained matrix $\boldsymbol{A}$ that represents the lead field matrix, with corresponding linear parameters $\boldsymbol{x}$ representing the sources in the surrogate model. Since we are focussed on data filtering and not source estimation, the use of a surrogate head model $\boldsymbol{A}$ is particularly appropriate. We will discuss a method for forming $\boldsymbol{A}$ in the Results Section. The combined total array collects data for n time slices to yield the spatio-temporal data matrix $\boldsymbol{D}$, modeled as

$$\boldsymbol{D} \cong \boldsymbol{AX}+\boldsymbol{N}. \tag{2}$$

We assume that both source $\boldsymbol{j}$ and noise $\boldsymbol{n}$ can be considered random vectors with known first and second order moments. We may specify or estimate the first and second moments of $\boldsymbol{n}$ with the assumption that $E(\boldsymbol{n}) \equiv \boldsymbol{0}$ and $E(\boldsymbol{nn}^T) \equiv \boldsymbol{C_n}$. Similarly, we may specify or estimate either $E(\boldsymbol{jj}^T) \equiv \boldsymbol{C_j}$ or $E(\boldsymbol{dd}^T) \equiv \boldsymbol{C_d}$, and both are assumed zero-mean for convenience here. We make the common assumption of independence between signal and noise, linking the two covariances into the model of *data covariance*,

$$\begin{aligned} \boldsymbol{C_d} &= (\boldsymbol{LC_jL}^T \\ &\quad +\boldsymbol{C_n}) \cong (\boldsymbol{AC_xA}^T+\boldsymbol{C_n}). \end{aligned} \tag{3}$$

We assume these matrices are decomposable and invertible as $\boldsymbol{C_n^{-1}}=\boldsymbol{W_n^{-T}W_n^{-1}}$, and $\boldsymbol{C_d^{-1}}=\boldsymbol{W_d^{-T}W_d^{-1}}$.

In many arrays, a specific physical subset of the array is designated as a *reference array*, typically placed at a distance from the head in order to reduce the possibility of recording neural activity. We will designate the other portion of the array closest to the head as the *primary array*.

2. METHOD

2.1. The Generalized Sidelobe Canceller

The high-dimensionality of the MEG array allows us to exploit the overconstrained head model $\boldsymbol{A}$. For the first step, because $\boldsymbol{A}$ is not of full row-rank, we can easily design its left null space operator, $\bar{\boldsymbol{U}}_A$. We apply this spatial "blocking" operator to the data in order to synthesize the *reference* time series,

$$\begin{aligned} \boldsymbol{D}_r &\equiv \overline{\boldsymbol{U}}_A^T \boldsymbol{D} \\ &= \overline{\boldsymbol{U}}_A^T (\boldsymbol{AX}+\boldsymbol{N}) = \overline{\boldsymbol{U}}_A^T \boldsymbol{N}. \end{aligned} \tag{4}$$

In other words, the result is a spatio-temporal *virtual* reference time series $\boldsymbol{D}_r$ devoid of any desired neural activity in the head model $\boldsymbol{A}$. This matrix does, however, contain a subset of the linear combinations of the noise sequences in the rows of $\boldsymbol{N}$.

In the second step, we find a linear fit of the *temporal* patterns between the reference time series and the original data, i.e., we model a linear relationship between the rows of $\boldsymbol{D}$ and $\boldsymbol{D}_r$ through a reference weighting matrix $\boldsymbol{W}_r$ as

$$\boldsymbol{D}^T = \boldsymbol{D}_r^T \boldsymbol{W}_r + \boldsymbol{E}, \tag{5}$$

where $\boldsymbol{E}$ is the error in this model. The weights are fit as an ordinary least-squares estimation between the synthetic reference and full array,

$$\begin{aligned} \boldsymbol{W}_r &= (\boldsymbol{D}_r^T)^{\dagger} \boldsymbol{D}^T \\ &= (\boldsymbol{D}_r \boldsymbol{D}_r^T)^{-1} \boldsymbol{D}_r \boldsymbol{D}^T. \end{aligned} \tag{6}$$

The result is an estimation of the noise-only sequences in the data, $\boldsymbol{W}_r^T \boldsymbol{D}_r$. To complete this step, we then filter the original data by removing this estimate,

$$\boldsymbol{F} = \boldsymbol{D} - \boldsymbol{W}_r^T \boldsymbol{D}_r = \boldsymbol{D} - \boldsymbol{D}\boldsymbol{D}_r^T (\boldsymbol{D}_r \boldsymbol{D}_r^T)^{-1} \boldsymbol{D}_r. \tag{7}$$

In this expression, we recognize the idempotent projection operator

$$\Pi \equiv \boldsymbol{D}_r^T (\boldsymbol{D}_r \boldsymbol{D}_r^T)^{-1} \boldsymbol{D}_r = \boldsymbol{V}\boldsymbol{V}^T, \tag{8}$$

where $\boldsymbol{V}$ is the orthonormal matrix spanning the space of $\boldsymbol{D}_r^T$. A compact expression for the GSC estimate is therefore

$$\boldsymbol{F} = \boldsymbol{D} - \boldsymbol{D}\Pi = \boldsymbol{D}\Pi^{\perp}. \tag{9}$$

Thus a virtual reference time series, void of any desired neural activity, has been used to create a subspace of the noise, and the data matrix has been temporally projected away from this space using the orthogonal complement projection operator $\pi^{\perp}$. Because the reference time series are *synthesized* using knowledge of the source model, this technique is termed a *generalized sidelobe canceller* (GSC), introduced by [3], first applied to electrophysiological data by [8], and further demonstrated for MEG in [5]. See [13] for an excellent overview of this and other beamformer techniques.

2.2. Other Methods

The GSC generates a *virtual* reference array void of neural activity, then adapts the data in this array to the primary array. In [14], we effectively generated (6) from a *physical* reference array known to be absolutely void of neural signals, yielding a *multiple sidelobe canceller* [13]. In most other physical reference arrays, however, the proximity of the sensors to the brain may allow neural activity to be recorded, and therefore direct application of (6) to such data could result in the cancellation of desired signals. Hence such systems generally provide the user with *non-adaptive reference weights*, presumably fitted from empty room data, i.e. (6) is

generated from noise priors. In *signal space projection* (SSP) [11], the dominant noise basis vectors are used to parse the data array into virtual primary and reference arrays, again yielding a form of (6) based on subspaces of noise priors. Finally, in *noise whitening*, we premultiply the data by $\boldsymbol{W}_n^{-1}$ to invertibly suppress the stronger noise spaces, yet allow strong signals to pass, as compared to SSP that truncates this space.

We next exploit the use of the overconstrained head model. The *generalized least-squares solution* for $\boldsymbol{X}$ is found as the *ordinary least-squares solution* of the *noise-whitened* data

$$\begin{aligned}\widehat{\boldsymbol{X}} &= (\boldsymbol{W}_n^{-1}\boldsymbol{A})^{\dagger}(\boldsymbol{W}_n^{-1}\boldsymbol{D}) \\ &= (\boldsymbol{A}^T\boldsymbol{C}_n^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{C}_n^{-1}\boldsymbol{D}.\end{aligned} \tag{10}$$

If (3) is true, i.e. $\boldsymbol{C_d} = (\boldsymbol{AC_xA^T} + \boldsymbol{C_n})$, then we may *identically* substitute the data covariance to yield

$$\begin{aligned}\widehat{\boldsymbol{X}} &= (\boldsymbol{W}_d^{-1}\boldsymbol{A})^{\dagger}(\boldsymbol{W}_d^{-1}\boldsymbol{D}) \\ &= (\boldsymbol{A}^T\boldsymbol{C}_d^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{C}_d^{-1}\boldsymbol{D},\end{aligned} \tag{11}$$

also known as the *linearly constrained minimum variance* [3], [12]. If the sample covariance is formed from $\boldsymbol{D}$ and used in (11), then one can show that $\boldsymbol{F} = \boldsymbol{A}\hat{\boldsymbol{X}}$ is *identically* (7). Finally, *signal space separation* (SSS) and its *temporal* extension (tSSS) [9], [10] use multipolar models of $\boldsymbol{A}$ and the space outside of $\boldsymbol{A}$ to fit the data, with tSSS followed by a partial adaptation of (6). Due to the limited space here, the proofs of these latter statements will be provided in a future publication.

3. RESULTS

We demonstrate the effectiveness of the proposed algorithm by applying it to human data acquired in the presence of substantial external and human artifacts. The data were acquired in a relatively lightly shielded room in an urban environment, such that the MEG magnetometers recorded substantial external disturbances. Additionally, the human subject in this example had lightly magnetized implants and dental work that created additional artifacts inside the shielded room.

The data were acquired on an MEG instrument comprising 102 magnetometers and 204 planar gradiometers. We intentionally restricted ourselves here to just the instrument's smaller magnetometer-only array, in order to demonstrate the applicability of this procedure to other relatively modest-sized arrays. The data were recorded at 1000 samples per second, with an analog high pass of 0.3 Hz and an analog low pass of 333 Hz. In post processing, the data were properly decimated to 200 samples per second for 40,600 samples total.

The subject's left and right median nerves were stimulated in a standard somatosensory evoked field protocol, for 266 repetitions for each hand over a total of about 200 seconds. The continuously recorded raw data are shown in Fig. 1. The data were dominated by large external magnetic disturbances.

Just prior to the human experiment, two minutes of empty room data were acquired to establish the noise covariance prior. An eigenanalysis of this data revealed a strong rank eight noise

space. We filtered the raw recordings by projecting away from this subspace, using the conventional SSP approach [11]. Separately, we whitened the data using the entire noise covariance matrix. The results of the whitened data are shown in Fig. 2, which were similar to the SSP-filtered data (not shown for brevity). The filtered data were now visibly dominated by a periodic signal of approximately four to five seconds wavelength, consistent with the respiration rate of the subject. Apparent movement artifacts were also visible at about 125, 175, and 200 seconds in the data.

With no additional artifact rejection of this data, we directly averaged the 266 trials of whitened data to yield the results shown in Fig. 3. Although an average response is becoming apparent, much artifact remains, which we conjecture is due to the strong respiration and movement artifacts.

We next built a relatively simple head model of a single sphere set 50 mm above the subject's coordinate origin, defined in the typical manner by connecting the subject's pre-auricular points and intersecting this line from the nasion. We synthesized a rough head shape of dipolar grid points by radially mapping the sensor centers inwards for several different depths, yielding 4,080 dipolar grid points spanning the interior volume of the helmet. For each point we generated the corresponding unconstrained dipole model, yielding a head model $102 \times 12{,}240$ columns [4], [6], [7]. A singular value decomposition of this matrix suggested a truncation of rank 70 was adequate, where the singular values were nearly four orders smaller than the dominant values. Thus the virtual reference array was of 32 dimensions, generated from remaining dimensions of the truncated head model.

The data were broken into seven second segments of 1400 data points, to yield 29 contiguous non-overlapping segments. This length was chosen based on experiences of [10] and [14], which suggest lengths of several seconds. Longer sequences in general are too dynamically changing for the adaptive filter to work as well, as we observed with this data set in separate tests not shown here. The GSC was applied otherwise unaltered to each segment of data, using (9). Processing time was only a few seconds on standard laptop. The resulting filtered data were now visibly whitened with little apparent artifact (not shown for brevity), and the same averaging was again applied to this filtered data to yield the results of Fig. 4. We now observe a dramatically smoother averaged data set, and the left and right median nerve responses are qualitatively visible in the figure. To confirm these results here as simply as possible, we separately ran MaxFilter (Elekta Neuromag) [2], [10], employing all 306 channels of data with a multipolar head model and temporal extensions (i.e. tSSS), to achieve nearly the same average waveforms (not shown for brevity).

4. Summary

Our demonstration here was kept intentionally simple, both for brevity and to highlight the primary parameters of the proposed adaptive filter. The head model was a simple spherical model with an approximate volume grid inside the helmet, and the model was truncated to rank 70 of 102 total dimensions. Future work will use more realistic head models and source locations, with more explicit justification of the truncation to build the virtual reference array. The data were parsed into non-overlapping segments of seven seconds; future work will further examine methods for selecting the segment length in a more continuous sequence of overlapping segments. As briefly discussed above, the GSC method can be shown in a natural framework of physical reference sensors, virtual reference sensors, noise whitening and maximum likelihood (generalized least-squares) estimation, each with similar assumptions of signal or noise spaces and signal-to-noise ratios, and the results of each of these methods can be compared.

For expository purposes, our example data here was the relatively well-known somatosensory evoked field response, locked to a known stimulus trigger. Of greater interest is the detection of interictal spikes, transient epileptic events that are often difficult to detect in raw recordings, even after whitening.

Acknowledgments

This work was supported in part by the National Institute of Biomedical Imaging and BioEngineering, grants R01-EB002010, R01-EB006572, and R01-EB009048, and the National Center for Research Resources, grant P41-RR14075.

References

1. Baillet S, Mosher JC, Leahy RM. Electromagnetic Brain Mapping. IEEE Signal Processing Magazine 2001;18(6):14–30.
2. Neuromag, Elekta. MaxFilter User's Guide, Software version 2.0. Oct. 2006
3. Griffiths L, Jim C. An alternative approach to linearly constrained adaptive beamforming. IEEE Trans Antennas Propagation 1982;30(1):27–34.
4. Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography - theory, instrumentation, and application to noninvasive studies of the working human brain. Rev Mod Phys 1993;65(2):413–497.
5. Mosher, JC.; Hämäläinen, M.; Leahy, RM.; Pantazis, D.; Volegov, P. Biomag. Vancouver, Canada: 2006 Aug 20–26. Generalized Sidelobe Canceller for Noise Reduction in MEG Arrays.
6. Mosher JC, Lewis PS, Leahy RM. Multiple dipole modeling and localization from spatio-temporal MEG data. IEEE Trans Bio Eng 1992;39(6):541–557.
7. Mosher JC, Leahy RM, Lewis PS. EEG and MEG: Forward solutions for inverse methods. IEEE Trans Bio Eng 1999;46(3):245–259.
8. Spencer, ME.; Leahy, RM.; Mosher, JC.; Lewis, PS. Adaptive filters for monitoring localized brain activity from surface potential time series. Proceedings IEEE. 26th Asilomar Conf. on Signals, Systems, and Computers; Pacific Grove, CA. October 1992; p. 156-160.
9. Taulu S, Kajola M. Presentation of electromagnetic multi-channel data: The signal space separation method. Applied Physics 2005;97:124905 1–10.
10. Taulu S, Simola J. Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. Phys Med Biol 2006;51(7):1759–1768. [PubMed: 16552102]
11. Uusitalo MA, Ilmoniemi RJ. Signal-space projection method for separating MEG or EEG into components. Med Biol Eng Computing 1997 Mar;35(2):135–40.
12. Van Veen BD, Van Drongelen W, Yuchtman M, Suzuki A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans Bio Eng 1997;44(9):867–80.
13. Van Veen BD, Buckley KM. Beamforming: A Versatile Approach to Spatial Filtering. IEEE ASSP Magazine 1988;5(2):4–24.
14. Volegov PL, Matlachov AN, Mosher JC, Espy MA, Kraus RH. Noise-free magnetoencephalography recordings of brain function. Phys Med Biol May;2004 49:2117–2128. [PubMed: 15214546]

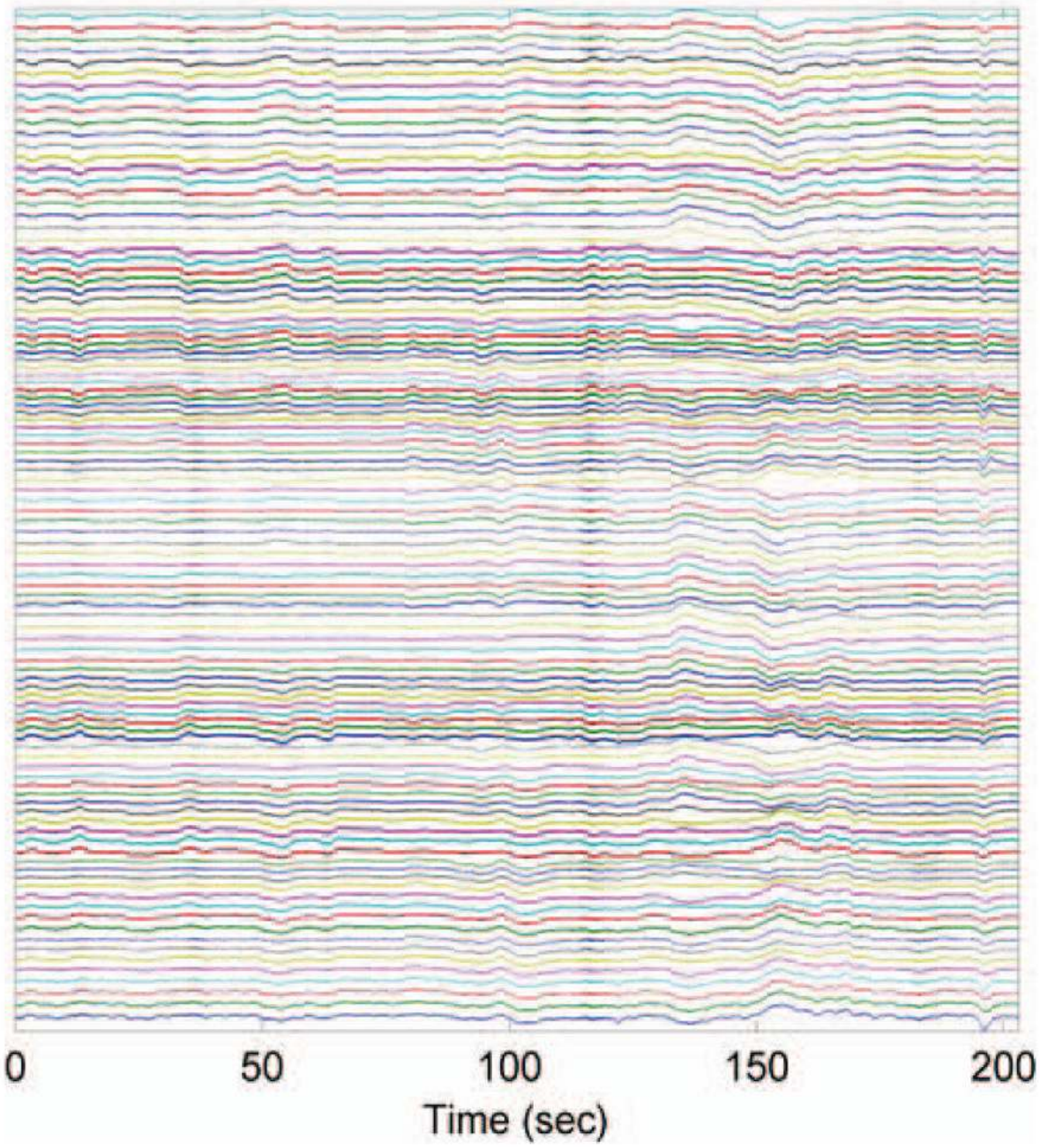

Fig. 1.
Raw recording from 102 magnetometers for 200 seconds, while 266 stimuli were applied separately to each left and right median nerves. Data were visibly dominated by external disturbances outside a lightly shielded room.

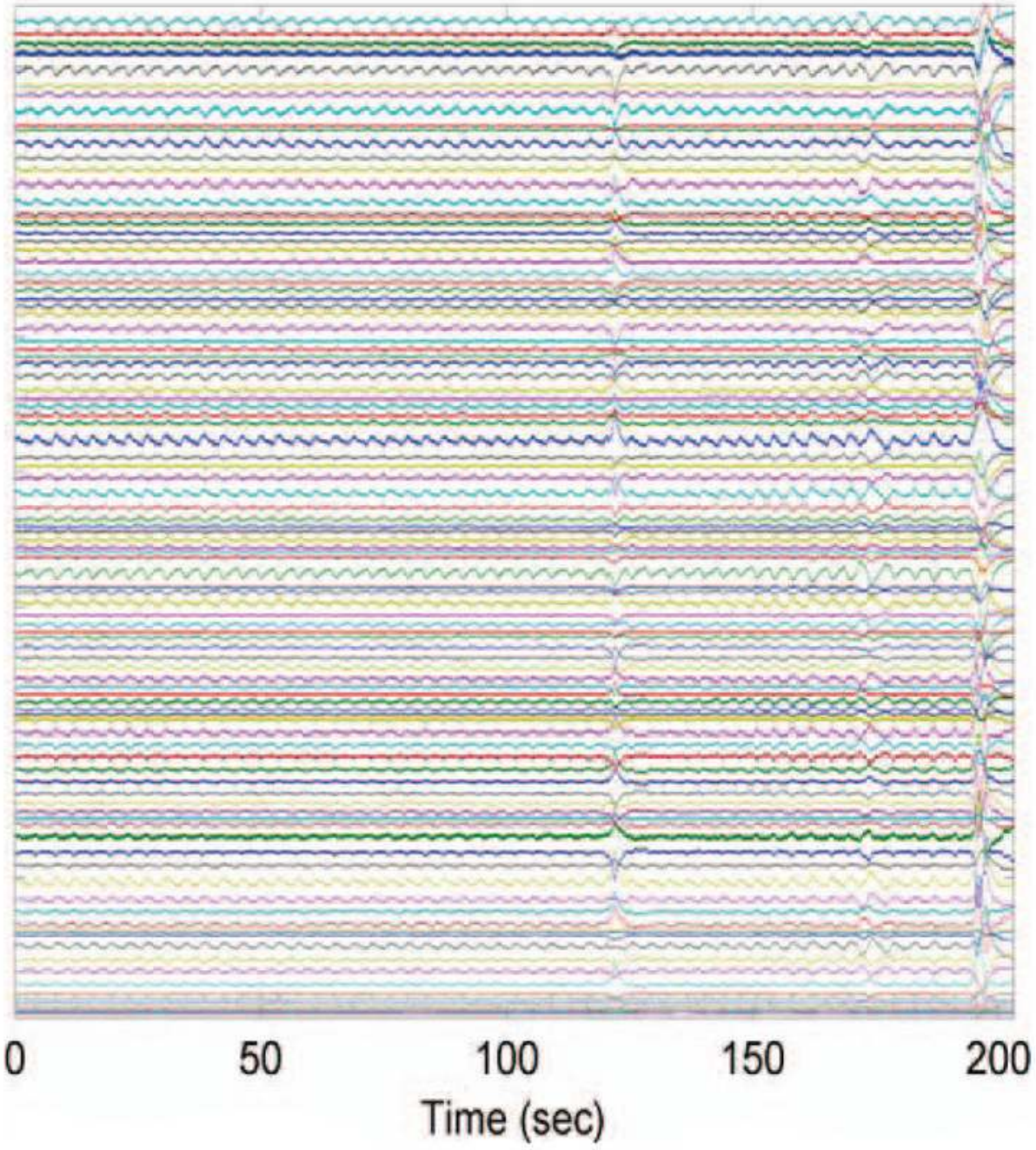

Fig. 2.
Using empty room data, the raw data were whitened. Data were now dominated by respiration artifacts, due to metal implants in subject. Transient movement artifacts by the subject were also now visible.

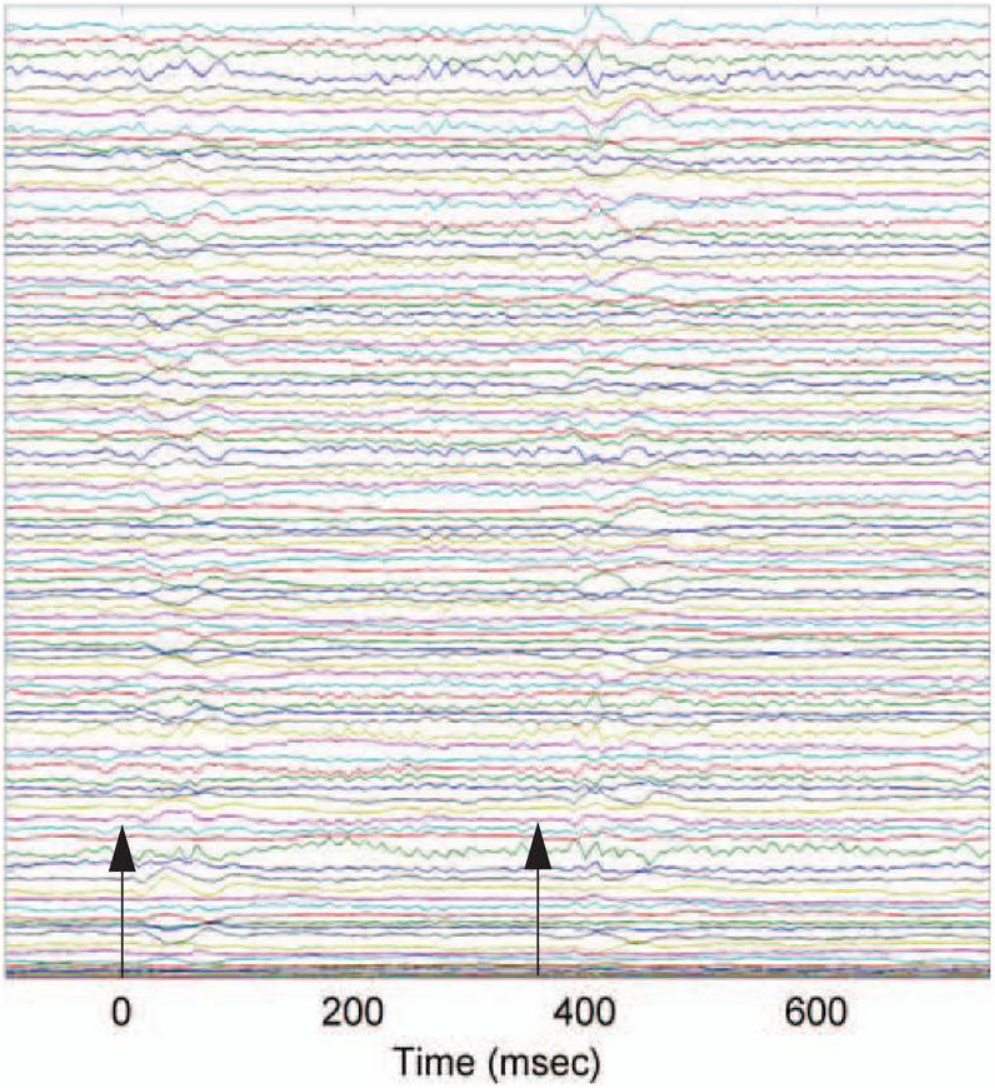

Fig. 3.

After whitening, average of the left and right median nerve data, from 266 stimuli. The left hand was stimulated at T=0, and the right hand was stimulated at T=375 ms. The responses are visible 25 – 50 ms after stimulus.

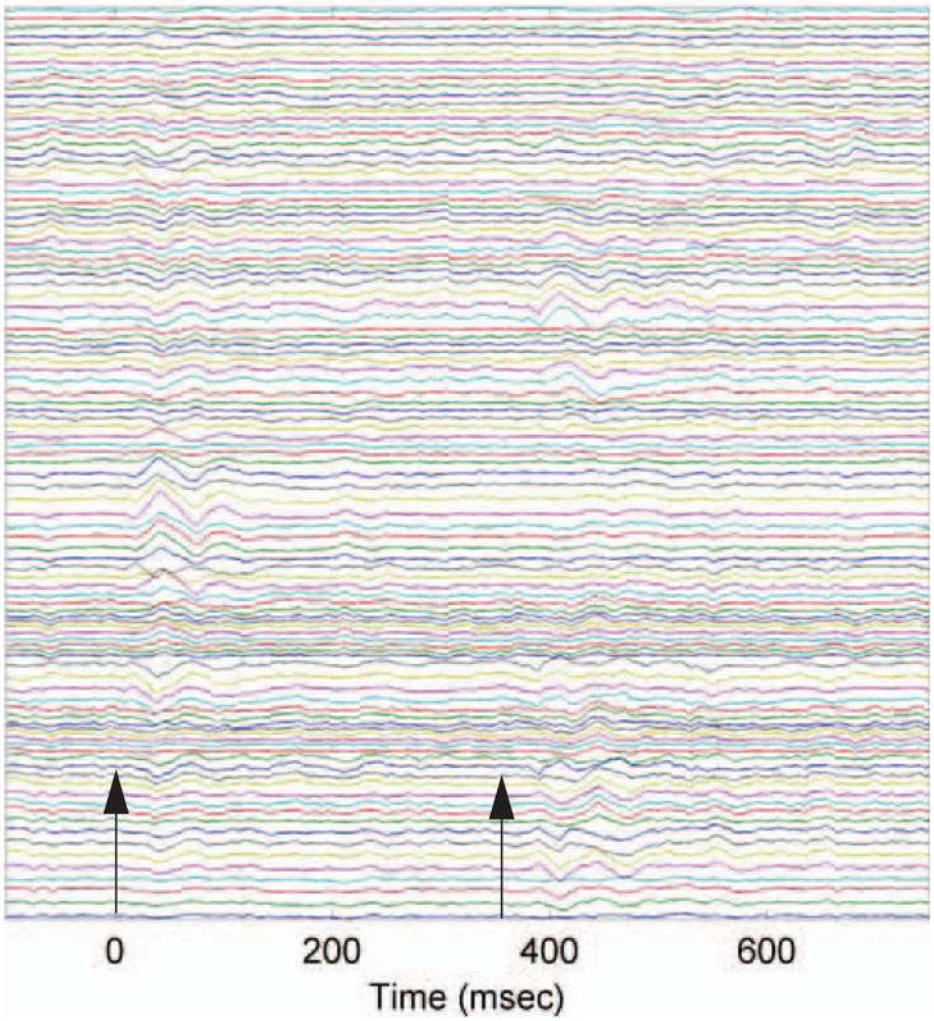

Fig. 4.
After filtering with the GSC algorithm, average of the same data. The two median nerve responses are now seem more clearly across the array, and the data are visibly smoother.