

# Linear spatial weighting for single trial detection in encephalography

Lucas Parra\*, Chris Alvino\*, Akaysha Tang†, Barak Pearlmutter†, Nick Yeung‡, Allen Osman§ and Paul Sajda||

\* *Vision Technologies Laboratory, Sarnoff Corporation, Princeton, NJ 08540, USA*

† *Department of Psychology and Department of Computer Science, University of New Mexico, Albuquerque, NM 87131, USA*

‡ *Department of Psychology, Princeton University, Princeton, NJ 08544, USA*

§ *Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA*

|| *Department of Biomedical Engineering, Columbia University, New York, NY 10027, USA*

## Abstract

Conventional electroencephalography (EEG) and magnetoencephalography (MEG) analysis often rely on averaging over multiple trials to extract statistically relevant differences between two or more experimental conditions. We demonstrate that by linearly integrating information over multiple spatially distributed sensors within a predefined time window, one can discriminate conditions on a trial-by-trial basis with high accuracy. We restrict ourselves to a linear integration as it allows the computation of a spatial distribution of the discriminating source activity. In the present set of experiments the resulting source activity distributions correspond to functional neuroanatomy consistent with the task (e.g. contralateral sensory-motor cortex and anterior cingulate).

## Introduction

Discriminating EEG/MEG activity on a single trial basis is essential for an efficient brain-computer interface (BCI), i.e. interpreting single trial brain activity for real-time communication. In the simplest case, where one wishes to communicate a binary decision, averaging corresponds to asking the same question over multiple trials and averaging the subject's binary decision responses. In order to obtain high-bandwidth communication it is desirable to do as little averaging (over time and/or trial) as possible. More generally, single trial analysis of brain activity is important in order to uncover the origin of response variability, and to study the variability of internal neurological processes.

Trial averaging is often used to increase the signal-to-interference (SIR) ratio, for example in analysis of event-related potentials (ERPs) [3]. With the large number of sensors in high density EEG and MEG an alternative approach is to integrate information over space rather than across trials. For example in blind source separation independent signals are extracted so that noise components and artifacts can be removed [8,13,14]. However, blind methods do not exploit the timing information of external events that is often available.

In the context of a BCI system, many methods have applied linear and non-linear classification to a set of features extracted from the EEG. For example autoregressive models have been used to extract features across a limited number of electrodes, with these features combined using either linear or non-linear classifiers to identify the activity from the time course of individual sensors [9]. Others have proposed to combine sensors in space by

computing maximum and minimum eigenvalues of the sensor covariance matrices to obtain a non-linear binary classification [10]. Though many of these methods show promising performance in terms of classifying covert (purely mental) processes, their neurological interpretation remains obscured.

We use conventional linear discrimination to compute the optimal spatial integration of a large array of sensors. We exploit timing information by discriminating and averaging within a short time window relative to a given external event. We demonstrate the method on three different electromagnetic brain imaging data sets: (1) predicting overt motor action (left/right button push) from 122 MEG sensors, (2) classifying covert or imagined motor activity (left/right taps) using 59 EEG sensors, (3) detecting decision errors in a binary discrimination task from 64 EEG sensors.

## Materials and Methods

*Linear Discrimination:* Denoting  $\mathbf{x}(t)$  as the  $M$  sensor values sampled at time instance  $t$ , we compute the spatial weighting coefficients  $\mathbf{v}$  such that

$$y(t) = \mathbf{v}^T \mathbf{x}(t) \quad (1)$$

is maximally discriminating between the times  $t$ , corresponding to two different experimental conditions. For example, in the prediction of explicit motor response experiments (described below) the times correspond to a number of samples prior to an overt button push. The samples corresponding to a left button push are to be discriminated from samples of a right button push. For each of  $N$  trials we may have  $T$  samples totaling  $NT$  training examples. We use conventional logistic regression [5] to find  $\mathbf{v}$ . After finding the optimal  $\mathbf{v}$  we average over the  $T$  dependent samples of the  $k$ th trial to obtain a more robust result,  $\bar{y}_k = T^{-1} \sum_{t \in T_k} y(t)$ , where  $T_k$  denotes the set of sample times corresponding to trial  $k$ . Receiver operating characteristic (ROC) analysis [12] is done using these single-trial short-time averaged discrimination activities ( $\bar{y}_k$ ). For visualization purposes, it is also useful to compute the trial averaged discrimination activities,  $\bar{y}_e(t) = N^{-1} \sum_{k \in N_e} y_k(t)$ , where  $N_e$  denotes the set of samples for event  $e$  (e.g. left or right button push) with time measured relative to some common reference across trials. The separation of the means together with their corresponding variances gives an indication of whether single-trial discrimination is plausible within the analysis window.

*Localization of Discriminating Sources:* In order to provide a functional neuroanatomical interpretation of the resultant spatial weighting, we treat  $y(t)$  as a source, and visualize the coupling coefficients of the source with the sensors. The strength of the coupling roughly indicates the closeness of the source to the sensor. The coupling  $\mathbf{a}$  is defined as the coefficients that multiply the putative source  $y(t)$  to give its additive contribution  $\mathbf{x}_y(t)$  to the sensor readings,  $\mathbf{x}_y(t) = \mathbf{a}y(t)$ . Unfortunately  $\mathbf{x}_y(t)$  is not observable in isolation, instead we observe,  $\mathbf{x}(t) = \mathbf{x}_y(t) + \mathbf{x}_{y'}(t)$ , where  $\mathbf{x}_{y'}(t)$  represents the activity that is not due to the discriminating source. If the contributions,  $\mathbf{x}_{y'}(t)$ , of other sources are uncorrelated with  $y(t)$  we obtain the coupling coefficients by the least-squares solution [7].

$$\mathbf{a} = \|\mathbf{y}\|^{-2} \mathbf{X}\mathbf{y}, \quad (2)$$

where the samples  $\mathbf{x}(t)$  for different  $t$  have been arranged as columns in the a matrix  $\mathbf{X}$ , and  $y(t)$  as a column vector  $\mathbf{y}$ . In general other sources are not guaranteed to be uncorrelated with

the discriminating source. Therefore, the "sensor projection"  $\mathbf{a}$  represents the coupling of all source activity that is correlated to the discriminating source  $y(t)$ . Our approach relies on the linearity of  $y(t)$  and the fact that different sources in EEG and MEG add linearly [1].

### *Datasets for Analysis*

*Prediction of motor response from MEG:* This data set was provided by AT and BP. Four subjects performed a visual-motor integration task. Subjects were simultaneously presented with two visual stimuli on a CRT. Subjects were instructed to push a left hand or right hand button, depending on which side a target stimulus was present. The subject was to discover the target by trial and error using auditory feedback. Each trial began with visual stimulus onset, followed by button push, followed by auditory feedback, indicating if the subject responded correctly. The interval between the motor-response and the next stimulus presentation was  $3.0 \pm 0.5$  s. Each subject performed 90 trials, which took approximately 10 minutes. MEG data was recorded using 122 sensor at a sampling rate of 300 Hz and high-pass filtered to remove DC drifts. Dipole fits were done using the Neuromag xfit tools, which assume a spherical head model to find a single equivalent current dipole.

*Classification of explicit and imagined motor response from EEG:* This dataset was provided by AO. Nine subjects performed a visual stimulus driven finger (L/R) tapping task. Subjects were asked to synchronize an explicit or imagined tap to the presentation of a brief temporally predictable signal. Subjects were trained until their explicit taps occurred consistently within 100 ms of the synchronization signal. After training, each subject received 10 blocks of trials. Each 72-trial block consisted of nine replications of the eight trial types (Explicit vs. Imagined x Left vs. Right vs. Both vs. No Tap) presented in a random order. Trials with noise due to eye blinks were not considered in the EEG analysis. The electromyogram (EMG) was recorded to detect muscle activity during imagined movements. The 59 EEG channels were sampled at 100 Hz and high-pass filtered to remove DC drift.

*Detection of decision errors from EEG:* This dataset was provided by NY. Seven subjects performed a visual target detection amongst distractors task. On each trial, subjects were presented with a stimulus for 100 ms. There were four possible stimuli, each consisting of a row of five arrows. Subjects were told to respond by pressing a key on the side indicated by the center arrow. They were to ignore the four flanking arrows. On half of the trials, the flanking arrows pointed in the same direction as the target (e.g. <<<<<<), on the other half the flankers pointed in the opposite direction (e.g. << >>>>). The interval between the motor-response and the next stimulus presentation was 1.5 s. Subjects performed 12 blocks of 68 trials each. The activity during an 100 ms interval prior to the response was used as the baseline. The sampling rate was 250 Hz. Following baselining, trials were manually edited to remove those with blinks, large eye movements, instrument artifacts and amplifier saturation.

## **Results and Discussion**

Single trial discrimination results are shown for the three different data sets and include trial averaged discriminating source activity  $\bar{y}_e(t)$ , sensor projections  $\mathbf{a}$ , and detection/prediction performance using single-trial, short-time averaged  $\bar{y}_k$ . Performance is reported using an ROC curve computed with a leave-one-out training and testing procedure [5]. Overall performance is quantified with  $A_z$  (the area under the ROC curve).

Figure 1 shows results for the dataset used to predict whether a subject will press a button with their left or right hand by analyzing the MEG signals in a window prior to the button

push. We use an analysis window 100 ms wide centered at 83 ms prior to the button event, which at 300 Hz corresponds to  $T=30$ . Figure 1 shows the results for one subject (AT). In the trial averaged results one can see significant separation of the means for left vs right button push within the analysis window. Single-trial discrimination is shown in the ROC curve, which for this subject shows good discriminability ( $A_z = 0.93$ ). Figure 1 also shows the sensor projection **a** and the location of a dipole-fit for this projection. When considered with respect to the motor-sensory homunculus, these results indicate that the discrimination source activity originates in the sensory-motor cortex corresponding to the left hand.

Figure 2 shows results for the second dataset, where the goal is to detect activity associated with purely imagined motor response, a situation more realistic for a BCI system. Subjects are trained to imagine a tap with the left or right index finger synchronized to a brief, temporally predictable signal. We selected a 0.8 s time window around the time where the task is to be performed. 90 left and 90 right trials were available to train the coefficients of the 59 EEG sensors. For the nine subjects we obtain a leave-one-out performance of  $A_z = 0.77 \pm 0.1$ . The results for the best performing subject is  $A_z = 0.90$ , shown in Figure 2. The sensor projection of the 59 EEG sensors shows a clear left-right polarization over the motor area. The results, across the nine subjects, for predicting explicit finger taps from a window 300 ms to 100 ms prior to the taps is  $A_z = 0.87 \pm 0.1$ . As shown in Figure 2, sensor projections of the discrimination vector for explicit motor response coincide with those for the imagined motor response. This is an important experimental finding supporting the development of an intuitive BCI system—signals arising from the cortical areas that encode an explicit movement can be also used for predicting the imagined movement.

Figure 3 shows the results for the experiments where the goal is to detect the Error Related Negativity (ERN) on a single trial basis. The ERN is a negative deflection in the EEG signals following perceived incorrect responses in a discrimination task [6]. The detection of error potentials has been proposed as a means of correcting communication errors in a BCI system [11]. The ERN has a fronto-central distribution, suggesting a source in the anterior cingulate [4]. It begins around the time of the incorrect response and lasts roughly 100 ms thereafter. We use this time window for detection. 40 to 80 error trials and 300 correct trials were used for training 64 coefficients. The sensor projection, shown in Figure 3 for one subject, is representative of the results obtained for other subjects and is consistent with the scalp topography of the ERN. The detection performance for this subject was  $A_z = 0.84$  and is to be compared to  $A_z = 0.63$  when detecting ERN from the center electrode where maximal activity is expected. ROC performance across all 7 subjects was  $A_z = 0.79 \pm 0.05$

Our results demonstrate the utility of linear analysis methods for discriminating between different events in single-trial, stimulus driven experimental paradigms using EEG and MEG. A particularly important aspect of our approach is that linearity enables the computation of sensor projections for the optimally discriminating weighting. This localization can be compared to the functional neuroanatomy, serving as a validation of the data driven linear methods. In all three cases presented, we find that indeed the activity distribution correlated with the source that optimizes single-trial discrimination localizes to a region that is consistent with the functional neuroanatomy. This is important, for instance in order to determine whether the discrimination model is capturing information directly related to the underlying task-dependent cortical activity, or is instead exploiting an indirect cortical response or other physiological signals correlated with the task (e.g. correlations with the stimulus, eye movements, etc.) . Localization of the discriminating source activity also enables one to determine the neuroanatomical correlations between different discrimination tasks, as was demonstrated for explicit and imagined motor responses in EEG. Finally, this

method is applicable to other encephalographic modalities with linear superposition of activity, such as near infrared imaging [2].

## References

- [1] Baillet, S., Mosher, J.C. & Leahy, R.M. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, **18**(6), 14-30, (2001).
- [2] Boas, D.A., et al. Imaging the body with diffuse optical tomography. *IEEE Signal Processing Magazine*, **18**(6), 57-75, (2001).
- [3] Coles, M. G. H. & Rugg, M.D. Event-related brain potentials: An introduction. In M. D. Rugg & M.G.H. Coles (Ed.), *Electrophysiology of Mind*. Oxford University Press (1995).
- [4] Dehaene, S., Posner, M., & Tucker, D. Localization of a neural system for error detection and compensation. *Psychological Science*, **5**, 303-305, (1994).
- [5] Duda, R., Hart, P. & Stork, D. *Pattern Classification*. John Wiley & Sons, 2<sup>nd</sup> ed., (2001).
- [6] Falkenstein, M., Hoorman, J., Christ, S. & Hohnsbein, J. ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology*, **51**:87-107, (2000).
- [7] Haykin, S. *Adaptive Filter Theory*. Englewood Cliffs, NJ, Prentice-Hall, (1996).
- [8] Makeig, S. Bell, A. Jung, T. & Sejnowski T. Independent component analysis of electroencephalographic data. In *Advances in Neural Information Processing Systems*, **8**, pages 145-151. MIT Press. (1996).
- [9] Pfurtscheller, G. & Neuper, C. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, **89**(7), 1123-1134, (2001).
- [10] Ramoser, H., Mueller-Gerking, J. & Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transaction on Rehabilitation Engeneering*, **8**(4):441-446 (2000).
- [11] Schalk, G., Wolpaw, J., McFarland, D., & Pfurtscheller, G. EEG-based communication: presence of an error potential. *Clinical Neurophysiology*, **111**:2138-2144, (2000).
- [12] Swets JA. ROC analysis applied to the evaluation of medical imaging techniques. *Investigative Radiology* **14**:109-121, (1979).
- [13] Tang, A., Pearlmuter, P., Phung, D. & Carter, S. Independent components of magnetoencephalography, Part I: Localization. *Neural Computation*, to appear (2002).
- [14] Vigario, R., Sarela, J., Jousmaki, V., Hamalainen, M. & E. Oja. Independent component approach to the analysis of EEG and MEG recordings. *IEEE Transactions on Biomedical Engineering*, **47**(5):589-593 (2000).

## Acknowledgements

This work was support in part by the Defense Advanced Research Project Agency, the National Foundation for Functional Brain Imaging and the National Institute of Health.

**Figure 1:** MEG left/right button push prediction. (A) Trial averages  $\bar{y}_e(t)$  (solid curves) and standard deviations (dotted curves) of discriminating component for left (red) and right (blue) button pushes. Time is given in seconds. Black vertical line at  $t = 0$  s indicates timing of button push. Green vertical lines indicate the discrimination window. (B) Sensor projections  $\mathbf{a}$  for discrimination vector. (C) ROC curve for left vs. right discrimination. Area under the curve  $A_z = 0.93$ . (D) Dipole-fit of  $\mathbf{a}$  overlaid on MRI image.

**Figure 2:** Discrimination of imagined left/right finger taps. (A) trial averages  $\bar{y}_e(t)$  (solid curves) and standard deviations (dotted curves) of discriminating component for left (red) and right (blue) imagined. Time is given in seconds. Black vertical solid line at  $t = 0$  s indicates timing of visual stimulus that defines the action to be performed (left or right imagine). The subjects are trained to execute the task at around  $t = 1:25$  s. Green vertical lines indicate the discrimination window. (B) Dorsal view of sensor projections  $\mathbf{a}$ . (C) ROC curve for left vs. right discrimination. (D) Sensor projection of discriminating source for explicit finger tap.

**Figure 3:** Detection of decision errors with EEG. (A) trial averages  $\bar{y}_e(t)$  (solid curves) and standard deviations (dotted curves) of discriminating component for correct and error trials. The negative deflection after a button push response at  $t = 0$  s is the ERN. Green vertical lines indicate the discrimination window. (B) Dorsal view of sensor projections  $\mathbf{a}$ . (C) ROC curve for error vs. correct trials. Solid curve corresponds to discrimination using Eq. (1) and dotted line to discrimination with center electrode (#12). (D) Sensor projection of discriminating source for explicit finger tap.



