

# Predicting Reaching Targets from Human EEG

Paul S. Hammon, Scott Makeig, Howard Poizner, Emanuel Todorov, Virginia R. de Sa

{phammon, smakeig, hpoizner}@ucsd.edu, {todorov, desa}@cogsci.ucsd.edu

University of California, San Diego

## Abstract

In this paper, we show that externally recorded electroencephalogram (EEG) signals contain sufficient information to decode target location during a reach (Experiment 1) and during the planning period before a reach (Experiment 2). We discuss the application of independent component analysis and dipole fitting for removing movement artifacts. With this technique we get similar classification accuracy for classifying EEG signals during a reach (Experiment 1) and during the planning period before a reach (Experiment 2). To the best of our knowledge, this is the first demonstration of decoding (planned) reach targets from EEG. These results lay the foundation for future EEG-based brain-computer interfaces (BCIs) based on decoding of planned reaches.

## Introduction

Brain-computer interfaces (BCIs) are being developed for a variety of applications ranging from assistive technologies for patients with motor disabilities to entertainment devices. Across the wide range of applications, all BCI systems share the same set of underlying components, which can be broken down into three main segments: brain signal acquisition, brain state decoding, and computer-mediated performance of a task. Some BCIs decode the brain state into a set of discrete classes such as yes/no commands, while other BCIs decode continuous data such as a reaching trajectory.

One goal of BCI research is to develop systems capable of decoding neural representations of natural movement planning and execution. The large number of degrees of freedom, high complexity, and speed of natural movement pose particular challenges to building BCI systems of this type. In order to deal with these constraints, researchers typically use arrays

of intracranial electrodes. There are several types of intracranial electrodes ranging from intracortical electrodes which measure the firing patterns of a small number of neurons (single unit recordings) to subdural or epidural electrodes measuring the local field potentials (electrocorticography, or ECoG) generated by dendritic currents in large numbers of nearby neurons. Relative to less-invasive recording methods, intracranial electrodes allow for higher spatial resolutions but suffer from increased costs and risks associated with surgical implantation and maintenance of the electrode array. To date, the vast majority of research in this area has focused on intracranial recordings in non-human primates; however, non-invasive recording techniques such as the electroencephalogram (EEG) offer benefits like improved safety and lower cost.

In this work, we analyze human EEG signals recorded from two different overt reaching tasks. We begin with data from a natural reaching task and show that reaching targets can be decoded from EEG. The promising results from this initial experiment led us to perform a second experiment based on a delayed reaching task. Results from this second experiment confirmed the promise from the first experiment that human EEG contains sufficient information to classify reaching targets.

## **Recording Brain Signals during Reaching**

There is a large body of BCI research using signals from intracranial electrodes in monkeys during reaching tasks. For example, in [16] Serruya et al. demonstrate that intracortical single unit recordings from a small number of primary motor cortex neurons can be used to reconstruct movement trajectories without extensive subject training. Another study demonstrates that recordings from the parietal reach region can be used to detect high-level movement goals [11]. Furthermore, the BCI system with the highest information transfer rate to date [15] is based on implanted electrodes in macaques. Recent work with ECoG data recorded during human reaching tasks [14] indicates that reaching data can be decoded from cortical activity at lower resolutions than in LFP or single unit recordings.

However, there is little work in this area using EEG data. There are two main reasons why EEG data is rarely used for reaching tasks. The first and perhaps most significant reason

is due to contamination from artifacts. EEG recordings pick up all electrical activity at the scalp and are therefore susceptible to pick up not only electrical signals generated by brain activity but also electrical activity generated by eye movements and contraction of head and neck muscles. Therefore, experiments using EEG are typically designed to minimize any movement. In order to use EEG data during a task which involves movement, it is important to isolate and remove artifacts (see Section for more details).

A second drawback of EEG is that it records the aggregate activity of large numbers of neurons. Thus, it is generally thought that techniques such as EEG do not extract detailed enough information to reconstruct complex movements.

Our goal is to show that—with appropriate artifact removal, signal processing, and machine learning—human EEG carries sufficient information about reach intention to decode reaching targets from the EEG signal during and even prior to planned movement.

## Experiment 1

EEG and 3D movement data were simultaneously recorded during a cued reaching task in a darkened room. EEG data were sampled at 256 Hz using a 256-channel Biosemi ActiveTwo system with sintered Ag-AgCl active electrodes. Movement data were recorded using a Northern Digital, Inc. Optrack Certus 3D tracking system. 3D locations of the hand, wrist, elbow, shoulder and the three target LEDs were sampled at 250 Hz.

Subjects sat in a comfortable chair and used their right index finger to reach for one of three target LEDs situated to the left (L), center (C), and right (R) of the subject’s sternum. The LED targets differed in height and distance from the sternum and were positioned such that each subject could reach them without resorting to a full arm extension. Two different starting postures of the arm were used: one with the right forearm horizontal and the upper arm vertical (H), and the other with the right arm fully flexed with the forearm vertical (V). An individual trial began with LED illumination with subjects instructed to make an immediate and natural reach to the target. The LED stayed on until the target was touched, or for 2500 ms in case of target miss. After touching the target, subjects were instructed to pause briefly and then return to the starting position. Target touch or time

out was followed by a random interval of 800-1200 ms before the next target appeared.

Starting pose was blocked, with subjects randomly assigned the starting pose for the first half (either H or V), with the remaining pose used for the second half (either V or H). There were 4 different sequences of 20 pseudo-random reaching targets each. Each sequence was run 4 times per starting pose condition. Thus there were a total of 320 trials per starting pose, and 640 trials overall.

## **Experiment 1 Data Pre-Processing**

The first step of data analysis was to reduce the dimensionality by cutting down the number of channels from 256 to 150. Channels with poor coupling or other problems were removed first, followed by channels located low on the head or neck, as these tend to have poor connections and large movement artifacts. Next channels were removed one at a time based on closest physical proximity to other electrodes until the desired number of channels (150 channels for S1, S3, and S4; 140 channels for S2) was met. Channel removal is a reasonable form of dimensionality reduction for high-density EEG because nearby channels are highly correlated, and thus little information is lost.

The EEG and movement data were aligned by re-sampling the movement data from 250 Hz to 256 Hz. The movement data were divided into two different segments: a planning (Plan) segment from LED illumination until movement was detected, and a movement (Move) segment consisting of forward movement to the target. Movement initiation was defined as the point at which the fingertip sensor reached 5% of the peak velocity during the initial acceleration phase. Termination of forward movement was defined as the point at which movement was at or near a minimum velocity and shifted direction, corresponding to the point in time at which the arm reversed direction and began returning to the starting position. Arm trajectories were visually inspected, and trials revealing data collection errors due to factors such as a subject starting a trial early or occluded LEDs were excluded from further analysis [13].

Movement data for the four subjects is summarized in Table 1 broken down by the three different reaching targets. Average times for both movement planning and execution are

listed in ms. An average of 612 trials was analyzed for each of the 4 subjects.

Table 1: Experiment 1 Subject Movement Data (Relative to LED Onset)

Subject	Right			Center			Left		
	Plan	Move	Trials	Plan	Move	Trials	Plan	Move	Trials
<b>S1</b>	362	675	203	342	629	207	346	664	202
<b>S2</b>	297	443	205	267	394	214	283	416	196
<b>S3</b>	252	523	206	242	472	214	240	485	212
<b>S4</b>	303	721	183	288	645	192	256	642	207

## EEG Artifact Detection and Removal

EEG data recorded during eye or other muscle movement typically contain movement artifacts. Artifacts associated with eye movement occur because the eye has an uneven charge distribution and therefore acts as an electric dipole. Eye blinks and other movements generate a varying electrical field that propagates throughout the head and is picked up by scalp electrodes. Muscles generate electrical activity when they contract. Muscle tension in scalp, face, and neck muscles generates signals which also propagate throughout the scalp adding electromyographic (EMG) activity to electrodes near the muscle insertion.

The standard approach to handling artifacts in EEG research is to avoid them by developing experiments which restrict movement as much as possible, and then discarding trials which contain movement artifacts. This approach, however, is impractical for natural reaching tasks, in which eye and muscle movement are generally unavoidable.

Because reach data is especially susceptible to contamination from EEG artifacts that could potentially lead to erroneous classification rates, we developed an especially conservative artifact removal procedure.

## Independent Component Analysis

The first step of our artifact removal approach is to run Independent Component Analysis (ICA) on the EEG data. ICA is a statistical technique which takes recordings from an array of sensors and determines a set of source signals which are maximally independent according to a specified measure of statistical independence. ICA posits a data model  $X = AS$ , where  $X$  are stacked row vectors of data recorded from individual sensors,  $A$  is a matrix of mixing

weights, and  $S$  are the stacked row vectors of statistically independent source vectors. Given only the data  $X$ , the ICA algorithm returns estimates for both  $A$  and  $S$ :  $\hat{S} = \hat{A}^{-1}X$ . The vectors in  $\hat{S}$  are termed “independent components” (ICs). The columns of  $\hat{A}^{-1}$  indicate how to construct an individual IC as a weighted combination of channels of  $X$ . Thus, we can visualize the distribution of an IC over the scalp by plotting the values from a column of  $\hat{A}^{-1}$  at each electrode location on the scalp to generate IC scalp maps (see Figure 1).

The ICA model assumes linear mixing of the sources as represented by the matrix  $A$ . Given a set of electrical sources in the brain, this linear mixing assumption holds because the net electrical potential between any two scalp electrodes is simply the superposition of the potentials resulting from each source.

There are a variety of ICA algorithms which primarily differ in the independence measure used. For artifact removal, we use InfoMax ICA [1, 9], which aims to minimize mutual information between sources by maximizing entropy. For ICA calculations and component visualization, we use EEGLAB [2].

### **ICA Dipole-based Artifact Removal**

Once ICA components have been recovered, we attempt to fit dipoles to the IC scalp maps using the DIPFIT plug-in for EEGLAB. This software package uses a four-shell spherical model (the brain, pial surface, skull, and scalp) and attempts to find either a single dipole or a pair of symmetric dipoles which best fit the IC scalp map. See [12] for details.

Each ICA component gets a dipole fit, but we only consider dipoles which are good fits ( $< 15\%$  residual variance between the IC scalp map and the fitting dipole scalp map). The location of the “good fit” dipoles are then checked against the head model, and only those dipoles which reside within the brain volume of the model are retained.

All ICA components which do not meet our dipole requirements are removed. The  $i$ th ICA component can be removed by first zeroing out column  $i$  of the mixing matrix  $\hat{A}$  to form a new matrix  $\hat{A}_i$ . Then the cleaned EEG data consisting of all but the removed component can be reconstructed as  $X_i = \hat{A}_i S$ .

The EEG data generated by back projecting only the “good fit” ICA components with

equivalent dipoles located within the brain will be referred to as “cleaned EEG.”

The scalp maps of the retained components for each of the four Experiment 1 subjects are shown in Figure 1.

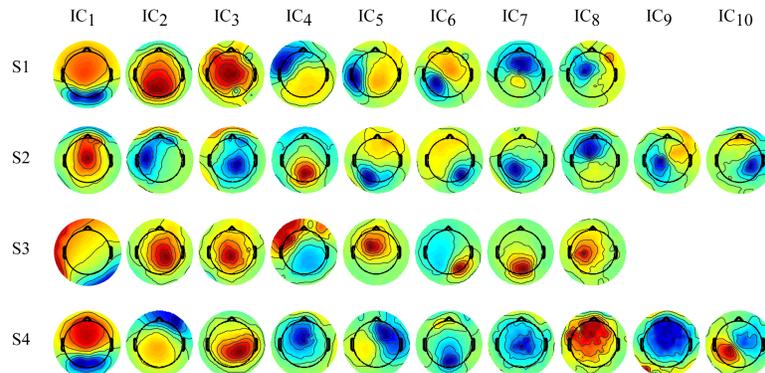


Figure 1: Scalp maps of the retained ICA components for all 4 Experiment 1 subjects. The hemispherical electrode locations are mapped to a flat disc, and thus electrodes further down on the head appear beyond the borders of the head outline.

## Feature Extraction

From the cleaned EEG data we extract a set of 8 different feature vectors which were useful for classification in previous BCI work [4].

### Feature 1: Autoregressive Modeling

An autoregressive (AR) system can be described by the difference equation  $y(t) = -\sum_k a_k y(t-k) + w(t)$ , where  $w(t)$  is a white-noise random process and  $a_k$  are the autoregressive coefficients. Thus, an AR system can be thought of as white noise passed through an all-pole filter. AR models give a decent first-order approximation of real EEG spectra but require far less data than high-resolution FFTs [17]. We use a 3rd order model and compute AR coefficients using the Burg method [10], parameters which we found to work well in previous analysis [4]. We apply this to each channel/component, resulting in a feature vector with 3 feature values per channel/component.

### Features 2: Power Estimates using a Filter Bank

A different approach to generating spectral estimates is to use a bank of filters to determine power estimates over frequency ranges corresponding to the individual filters. This approach

has added flexibility over direct FFT approaches and allows for shorter filters where lower resolution is required. We create a bank of 9th order finite impulse response (FIR) filters over the following frequency bands: delta (1-3 Hz), theta (4-7 Hz), alpha/mu (8-13 Hz), beta (14-25 Hz), and gamma (26-40 Hz). To generate the power estimate over a single frequency band, we filter the data, perform a point-wise squaring of the filtered signal, and finally compute the mean. This yields 5 feature values per channel/component.

### **Feature 3: Cleaned EEG Data**

The cleaned EEG data is the starting point for all features and thus has all (conservatively estimated) the information about the user’s brain state; however, the dimensionality of the data is too high. In order to reduce dimensionality, we re-sample the data to a fixed length of 10 samples per channel/component.

### **Feature 4: Discrete Wavelet Transform**

The discrete wavelet transform (DWT) offers a middle ground between frequency-based and time-based representations. We include a Symlet-based three-level DWT decomposition, which is applied to the data and then re-sampled to a total of 10 samples per channel/component.

### **Features 5 - 8: ICA Transforms**

Features 5 - 8 are generated by performing an ICA decomposition of the cleaned EEG data and then applying features 1 - 4, respectively. Because ICA is a data-dependent transform, this ICA decomposition is built into our classification cross-validation loops and trained only on the current training segment. We did not use our ICA components computed during artifact removal because they used all the data and we want to predict performance on future data not in our training dataset even when that data is not available during the construction of the classifier. (ICA on the whole dataset was used for the artifact removal because in that case we want to remove artifacts which would not be present in a movement-impaired target population and which might be artificially inflating our results.) Thus, ICA is run on each cross-validation loop, and computation speed is important. For performing ICA inside of our classification cross-validation loops, we use FastICA [5] instead of InfoMax

because the FastICA decomposition can be performed much more rapidly [6]. Prior to ICA decomposition, PCA is used to reduce dimensionality to 25.

## Classification Approach

Our classification approach involves creating individual classifiers based on each of the 8 different feature vectors, and then combining these individual classifiers into a meta-classifier [4]. Classification takes place in a nested cross-validation scheme. An outer 4-fold cross-validation loop segments the data into training (75%) and test (25%) sets. This outer loop runs 4 times such that each sample is a member of the test set exactly once. Eight individual classifiers are trained inside this loop. Within each outer cross validation loop, there is a 2-fold inner cross-validation loop in which the (outer) training data is once again segmented into (inner) training (50%) and test (50%) data. This inner cross validation loop is used to select classifier regularization parameters (described later). The individual classifiers are combined into a meta-classifier which is then tested on the test data. This procedure is outlined below.

1. outer cross-validation loop  $i = 1 : 4$ 
  - (a) loop over features  $j = 1 : 8$ 
    - i. extract feature  $j$
    - ii. inner cross-validation loop  $k = 1 : 2$ 
      - A. loop over regularization parameters  $m = 1 : 7$
    - iii. select best regularization parameter  $b$
    - iv. train classifier on feature  $j$  with regularization parameter  $b$
  - (b) combine individual classifiers to create meta-classifier  $i$
  - (c) compute classification rate on test data  $i$
2. compute average cross-validation classification rate

## Multinomial Logistic Classification

There are a variety of different classification approaches currently in use with BCI data. It is becoming increasingly clear that many different classifiers perform well given a sufficiently rich feature set [4]. This particular application has high-dimensional multi-class data. Thus,

we would like to use a classifier with regularization to avoid over-fitting and native multi-class support to avoid training multiple one-versus-rest classifiers. To address these two characteristics, we use sparse multinomial logistic regression [8].

The multinomial logistic regression classifier is:

$$p(y^{(i)} = 1 | \mathbf{x}; \mathbf{w}) = \frac{\exp(\mathbf{w}^{(i)T} \mathbf{x})}{\sum_{j=1}^m \exp(\mathbf{w}^{(j)T} \mathbf{x})}, \quad (1)$$

where  $i$  indicates the class number,  $y^{(i)}$  is 1 for class  $i$  and 0 otherwise,  $\mathbf{x}$  is the feature vector,  $\mathbf{w}^{(i)}$  is the weight vector for class  $i$ , and  $m$  is the number of classes.

The regularization comes from placing a sparsity-promoting Laplacian prior on the weight vectors  $\mathbf{w}^{(i)}$ :

$$p(\mathbf{w}^{(i)}) \propto \exp(-\lambda \|\mathbf{w}^{(i)}\|_1), \quad (2)$$

where  $\lambda$  is a regularization parameter and  $\|\cdot\|_1$  denotes the  $l_1$  norm. We use the inner cross-validation loop to select the best value of  $\lambda$  from  $\{0, 10^{-6}, 10^{-4}, 10^{-2}, 10^0, 10^2, 10^4\}$ .

### Combining Classifiers

Combining multiple classifiers has the potential of reducing error rates both by reducing variance and by combining possibly independent types of information [3, 4]. We combine the output of the 8 individual classifiers using an average meta-classifier: average the predicted probabilities for each class across all individual classifiers, and the class with the highest average probability is selected as the predicted class output. The average meta-classifier is robust to individual classifiers with poor probability estimates [7], and it has no parameters, which avoids extra steps for tuning the meta-classifier parameters.

## Experiment 1 Classification Results and Feature Analysis

We built classifiers to investigate two different aspects of the combined movement and EEG data: two-class classifiers to determine starting pose collapsed across target end points,

and three-class classifiers to determine target end point collapsed across starting positions. We built separate classifiers for each of the two different data segments, Plan and Move. Average cross-validation results are summarized in Table 2.

Table 2: Classification Error Rates ( $\pm$  Standard Deviation) for Target Reaching

Subject	Target End Point (3-class)		Starting Pose (2-class)	
	Plan	Move	Plan	Move
<b>S1</b>	0.502 $\pm$ 0.0705	0.459 $\pm$ 0.0415	0.0572 $\pm$ 0.0222	0.0147 $\pm$ 0.00327
<b>S2</b>	0.369 $\pm$ 0.0421	0.380 $\pm$ 0.0314	0.163 $\pm$ 0.0325	0.125 $\pm$ 0.0243
<b>S3</b>	0.342 $\pm$ 0.0301	0.263 $\pm$ 0.0285	0.0158 $\pm$ 0.0121	0.0142 $\pm$ 0.0108
<b>S4</b>	0.433 $\pm$ 0.0448	0.378 $\pm$ 0.0361	0.0533 $\pm$ 0.0249	0.0275 $\pm$ 0.0110

Reaching target error rates were somewhat high for the pre-movement interval (ranging from 34% to 50% error) and generally lower for the mid-movement interval (ranging from 26% to 46% error). While these error rates may seem somewhat high, they are all well below chance (67%), and half of the subjects tested had error rates in the 30% error range. These relatively high error rates may be at least partially due to the conservative artifact rejection scheme that reduced the rank of the EEG data from 150 (140 for S2) down to 8 - 10 ICA components.

It is worth noting that the starting pose is more discriminable than the three different reaching targets. This is not too surprising, as these two poses differ more than any pair of poses from reaches to different targets. However, the error rates on the order of 1% are somewhat suspicious. One possibility is that the high classification rates are at least partially due to the nature of the experimental design, in which all trials from a given starting pose were run as a block. Thus, starting pose could act as a proxy for first vs. second half of the experiment.

To better understand what specific aspects of brain activity were most important for target reach classification, we investigated the individual classifiers. One way to do this is to compare the absolute value of the classifier weights, which indicates the relative importance of the corresponding features. To compare feature weights across subjects, we constructed a reduced set of 69 channels which matched locations across all 4 subjects, scaled each feature vector to have the same average length, and only analyzed the channel-based features 1a -

4a, which could be lined up across subjects.

To visualize the feature weights, we took the feature weights (from the Plan stage) and averaged across cross-validation folds and subjects. For each feature, this gives 3 weight vectors (one per class) which we reduced to a single vector by computing the variance across the three different classes. These values were then mapped onto 2-D head images using a fixed color map range for each feature. Results are shown in Figure 2. Each row represents a complete feature vector, with an average scalp map for each frequency/time point.

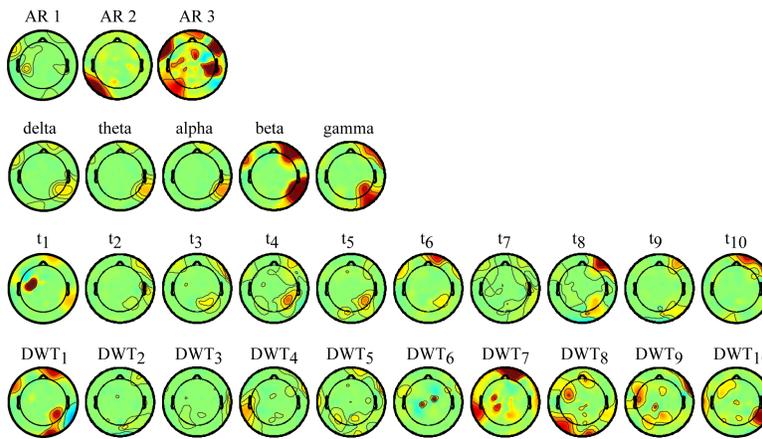


Figure 2: Scalp maps of feature importance averaged across all 4 Experiment 1 subjects.

The locations of the largest feature weights vary substantially in position across the 4 subjects and the different feature vectors. One feature of note is the large weights over left sensorimotor areas ( $t_1$ ,  $AR_1$ ), which lie near the hand area of primary motor cortex contralateral to the reaching hand. Some features show large weights over frontal areas (beta, gamma,  $DWT_7$ ). These could be due to frontal attention processes, but it is also possible that these are due to remnant eye artifacts. The occipital and temporal weights (beta, gamma,  $DWT_1$ ,  $DWT_{10}$ ) could be related to visual processing, but lie in an area which could also suggest neck muscle artifacts. All features showed large differences across subjects, reinforcing the notion that tailoring classifiers to each individual is crucial for good BCI performance.

## Experiment 2

The results from Experiment 1 are encouraging however, while our artifact reject scheme should remove most eye and muscle artifacts, it is possible that some artifact information was not successfully rejected and affected the classification results. Thus, we developed and ran a second experiment in which we had greater control over potential EEG artifacts.

Our second experiment is a center-out delayed reaching task modeled after the delayed reaching task employed by Santhanam et al. in [15]. In our variation, the subject sits back in a comfortable chair positioned within easy reaching distance of a 19 in. LCD touch screen (Elo TouchSystems Model 1925L). EEG data were recorded at 256 Hz using a 64-channel Biosemi ActiveTwo system with sintered Ag-AgCl active electrodes. Two subjects were run in this new paradigm, one who also participated in Experiment 1 (S2) and one new subject (S5).

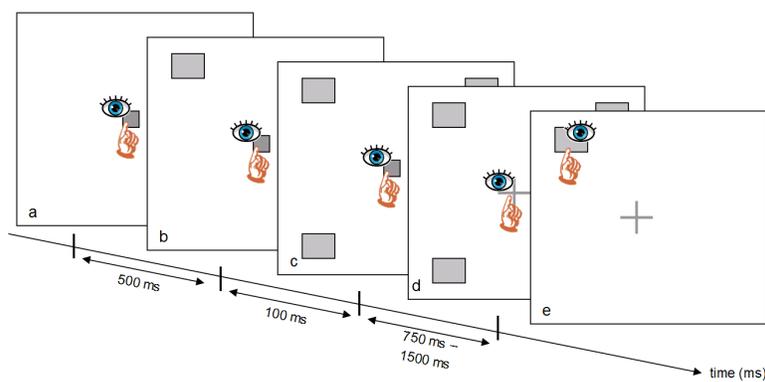


Figure 3: A single trial from Experiment 2.

A single trial of the experiment is diagramed in Figure 3. Subjects are instructed to touch, hold, and fixate on a central target (Figure 3a). After a short delay, a reach target briefly appears (100 ms) in one of the four corners of the screen (Figure 3b). All four possible targets are then shown and the experiment begins a variable delay phase (randomly selected between 750 ms and 1500 ms) during which the subject maintains touch and fixation of the central fixation (Figure 3c). After the variable delay, a “go” cue is indicated by replacing the central rectangular target with a cross (Figure 3d). At this point the subject is instructed to perform a natural reach to the (remembered) correct target. In order to keep the reach

as natural as possible, the subjects were not instructed to maintain central fixation during the actual reach. Touching the correct target before reach timeout (3 s) results in a “ding,” while an incorrect reach or timeout is signaled by a “buzz” (Figure 3e).

A total of 800 trials were run for each subject with short breaks every 25 trials. The 4 targets were presented in pseudorandom order with a total of 200 trials for each target.

## Experiment 2 Data Pre-Processing

The EEG data in this experiment were processed in exactly the same manner as those from Experiment 1, with only a few minor differences noted below. We used exactly the same algorithms for artifact removal, feature extraction, and classification. Scalp maps of the retained components for both Experiment 2 subjects are shown in Figure 4. The differences noted below stem primarily from the different number of channels and different time course of this experiment.

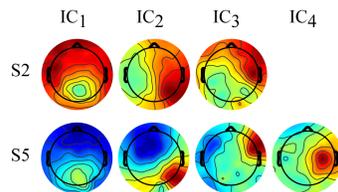


Figure 4: Scalp maps of the retained ICA components for both Experiment 2 subjects.

For Experiment 2, we only analyzed the first 500 ms of the delay period. During this period the visual stimulus is identical for all trials and visual fixation and arm position are held fixed. The intention was to decrease as much as possible the creation of any artifacts associated with eye and arm movements and systematic posture changes. Trials with incorrect target selection or reach timeout were excluded. The horizontal and vertical electrooculogram (HEOG, and VEOG, respectively) were visually inspected, and any trials showing blink or eye movements were excluded. Before classifying the data, we randomly selected a balanced number of trials from each class. The numbers of trials analyzed is summarized in Table 3.

## Experiment 2 Results and Feature Analysis

Three different classifiers were constructed for each subject: a 4-class classifier for reaching target, and 2-class classifiers dividing the target space into left vs. right reaches and top vs. bottom reaches. Average cross-validation results are summarized in Table 3.

Both subjects had similar error rates in 4-class target decoding, with both well below chance (75%). The two-class data showed an interesting pattern where classifiers for both subjects were substantially better at discriminating left from right targets than at discriminating top from bottom targets. These results are encouraging and further support our finding that reach targets can be successfully decoded from human EEG.

Table 3: Experiment 2 Subject Data and Results

Subject	Trials	Classification Error Rates ( $\pm$ Standard Deviation)		
		4-Class	Left v. Right	Top v. Bottom
<b>S2</b>	544	0.439 $\pm$ 0.0645	0.265 $\pm$ 0.0708	0.313 $\pm$ 0.0272
<b>S5</b>	536	0.414 $\pm$ 0.0325	0.142 $\pm$ 0.0305	0.326 $\pm$ 0.0347

We analyzed the features from Experiment 2 exactly the same way as in Experiment 1. The resultant feature maps are shown in Figure 5. Once again, weights vary substantially across subjects and features. Several features showed large weights over left and right sensorimotor areas (delta, alpha,  $DWT_{3,4,5,8}$ ) which would be consistent with reach planning. Note that there is little activity over occipital/parietal regions (with the possible exception of gamma), suggesting that any neck muscle artifacts had little impact on classification. There are some large frontal weights ( $t_1$ ,  $DWT_{1,6,10}$ ) which might correspond to frontal attention processes. These could be eye artifacts, but this is less likely than in Experiment 1 because all trials with apparent eye movements were removed.

### Error Patterns

One particularly interesting question about the EEG/movement data is whether or not it contains an underlying structure. One way such a structure might manifest itself is through structured errors. For example, in Experiment 1 one might expect that misclassified L trials are more likely to be (incorrectly) classified as C trials than R trials. To test this, we

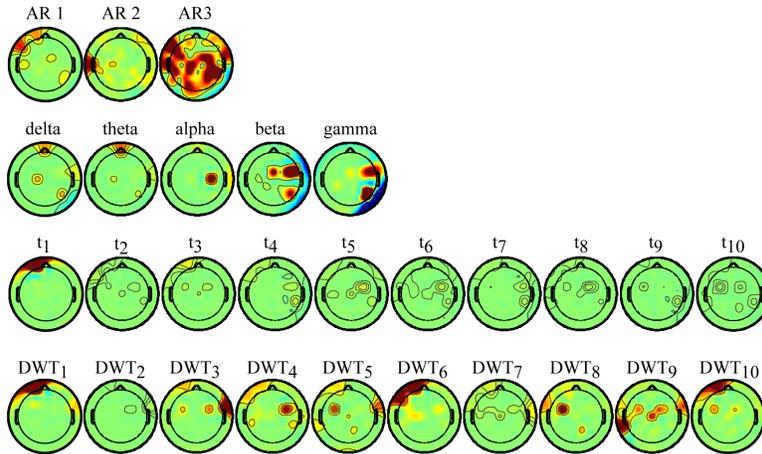


Figure 5: Scalp maps of feature importance averaged across Experiment 2 subjects.

constructed confusion matrices for target endpoints for Experiments 1 and 2.

The average confusion matrix for Experiment 1 is shown in Figure 6a, and the confusion matrix for Experiment 2 is shown in Figure 6b. Rows indicate the correct class label, and columns indicate the predicted class label. Thus, each cell shows what fraction of examples from a given class was predicted to belong to that class. For example, in Figure 6a row 1, column 2 (correct: L, predicted: C) shows the fraction of L trials which the classifier predicted to belong to class C in Experiment 1.

The structure of the Experiment 1 confusion matrix reveals that trials are most likely to be correctly classified, less likely to be misclassified one target away, and least likely to be misclassified 2 targets away. The Experiment 2 confusion matrix has a more complex structure suggesting that targets are most likely to be misclassified as nearby targets, but targets on the left (3, 4) are unlikely to be misclassified as targets on the right (1, 2), and vice-versa. These results suggest that there is an underlying structure to the target space, whereby targets nearby in physical space are also nearby in EEG feature space, but this target space may have additional complexities as it does in Experiment 2.

## Conclusions

We have successfully classified reach targets from human EEG in natural and delayed reaching tasks. We believe that our conservative artifact removal scheme should be sufficient to

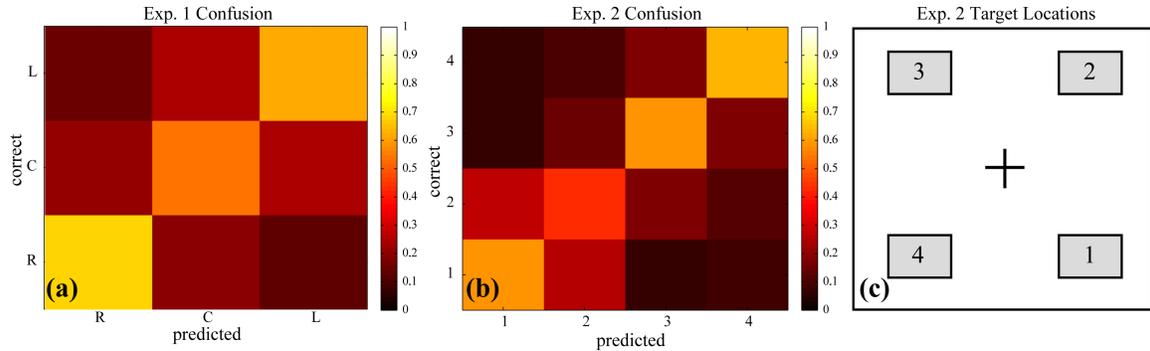


Figure 6: Average confusion matrices for Experiment 1 (a) and Experiment 2 (b). These matrices illustrate what fraction of the data was classified as each of the possible predicted class labels. Lighter shades indicate larger fractions. Target labels for Experiment 2 are shown in (c).

remove most movement-related EEG artifacts. This conjecture is supported by good classification performance on our delayed reaching task in Experiment 2, suggesting that our classification results on the natural reaching task in Experiment 1 are due to brain activity and not merely motion artifacts. The fact that the cleaned EEG data, especially in the planned reaching task, contained sufficient information to interpret the planned reach targets indicates that: 1) EEG is a viable recording technique for studying brain dynamics (without movement artifacts) and 2) EEG-based BCIs in people without motor control may be able to read their brain waves and act on their desired reach commands.

## Acknowledgements

We thank Julie Onton and Miro Enev for assistance with movement decoding and data analysis, and Andrey Vankov for assistance with data acquisition.

This work was supported in part by NSF Grant DGE-0333451 and NSF CAREER Award 0133996, a grant from the Kavli Brain-Mind Institute at UCSD, NSF Center grant “The Temporal Dynamics of Learning”, and NIH grant # 2 R01 NS36449.

## References

- [1] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159, 1995.
- [2] A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, 134:9–21, 2004.

- [3] Guido Dornhege, Benjamin Blankertz, Gabriel Curio, and Klaus-Robert Müller. Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms. *IEEE Trans. Biomed. Eng.*, 51:993–1002, 2004.
- [4] Paul S. Hammon and Virginia R. de Sa. Pre-processing and meta-classification for brain-computer interfaces. *IEEE Trans. Biomed. Eng.*, 54(3):518–525, 2007.
- [5] Jarmo Hurri, Hugo Gävert, Jaakko Särelä, and Aapo Hyvärinen. FastICA package for matlab. Available online at <http://www.cis.hut.fi/projects/ica/fastica/code/dlcode.shtml>, 2005.
- [6] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. Wiley, New York, 2001.
- [7] Josef Kittler, Mohamad Hatef, Robert P. W. Duin, and Jiri Matas. On combining classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20:226–239, 1998.
- [8] Balaji Krishnapuram, Lawrence Carin, Mário A.T. Figueiredo, and Alaxandar J. Hartemink. Sparse multinomial logistic regression: fast algorithms and generalization bounds. *IEEE Trans. on Pattern Analysis and Machine Learning*, 27(6):957–968, 2005.
- [9] S. Makeig, Anthony J. Bell, Tzyy-Ping Jung, and Terrence J. Sejnowski. Independent component analysis of electroencephalographic data. In D. Touretzky, M. Mozer, and M. Hasselmo, editors, *Advances in Neural Information Processing Systems 8*, pages 145–151, 1996.
- [10] S. Lawrence Marple. *Digital Spectral Analysis: With Applications*. Prentice Hall, 1987.
- [11] S. Musallam, B. D. Corneil, B. Greger, H. Scherberger, and R. A. Andersen. Cognitive control signals for neural prosthetics. *Science*, 305:258–262, 2004.
- [12] Julie Onton and Scott Makeig. Information-based modeling of event-related brain dynamics. *Prog Brain Res.*, 159:99–120, 2006.
- [13] H. Poizner, L. Mack, M. Verfaellie, L. J. G. Rothi, and K. M. Heilman. Three-dimensional computergraphic analysis of apraxia. *Brain*, 113:85–101, 1990.
- [14] Justin C. Sanchez, Aysegül Gunduz, Paul R. Carney, and Jose C. Principe. Extraction and localization of mesoscopic motor control signals for human ECoG neuroprosthetics. *J Neurosci Methods*, doi:10.1016/j.jneumeth.2007.04.019, 2007.
- [15] Gopal Santhanam, Stephen I. Ryu, Byron M. Yu, Afsheen Afshar, and Krishna V. Shenoy. A high-performance brain-computer interface. *Nature*, 442:195–198, July 2006.
- [16] M. D. Serruya, N. G. Hatsopoulos, L. Paninski, M. R. Fellows, and J. Donoghue. Instant neural control of a movement signal. *Nature*, 416:141–142, 2002.
- [17] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, and Theresa M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113:767–791, 2002.