Finger Movement Classification for an Electrocorticographic BCI

Pradeep Shenoy Kai J. Miller Jeffrey G. Ojemann Rajesh P.N. Rao
Dept. of Computer Science and Engineering
University of Washington
{pshenoy,kai,rao}@cs.washington.edu, jeff.ojemann@seattlechildrens.org

Abstract— **We study the problem of distinguishing between individual finger movements of one hand using electrocorticographic (ECOG) signals. In previous work, we have shown that ECOG signals have high predictive accuracy and spatial resolution for classifying hand versus tongue movements. In this paper, we significantly extend this paradigm by studying the first 5-class classification problem for ECOG, and show that an average 5-class error of 23% across 6 subjects is possible using as little as 10min of training data. In addition to opening up possibilities for higher-bandwidth brain-computer interfaces, the use of finger movements for control may yield a more intuitive mapping from ECOG signals to control of a prosthetic. Although this study uses real movements, our results provide the foundation for understanding ECOG signal changes during finger movement.**

I. INTRODUCTION

A vast body of recent work [1], [2], [3], [4], [5] deals with the problem of translating brain signals into control of a computer or robotic device. Such a control scheme, known as a *Brain-Computer Interface* or BCI, could serve as a communication mechanism for paralyzed or otherwise incapacitated individuals.

A popular paradigm for BCI is the classification of noninvasive electroencephalographic (EEG) signals recorded during real or imagined movements of various body parts (see e.g., [2]). Due to the noisy artifact-prone nature of EEG, and the lack of spatial resolution (measured activity is usually smeared over a large area of the scalp), only 2-3 classes of motor activity are used in typical BCI scenarios. Thus, a principal challenge for BCI is to discover more control signals [6], [7], [8] and analysis methods that can reliably distinguish between these control signals.

Recent work [3] has shown that electrocorticographic signals can be used to very quickly achieve control over a cursor, and that ECOG signals contain information in high-frequency spectral bands not available in EEG. The high-frequency band power shows distinct spatial activation patterns for different body parts [9]. Our recent work [10] demonstrated that the same set of spatial and spectral features can be used to classify ECOG signals from both real and imagined hand and tongue movements across a number of subjects. ECOG signals from real movements can be classified more accurately than EEG signals, using substantially lesser amounts of training data. Thus, a growing body of work indicates that ECOG signals may be a promising signal source for BCI.

In this paper, we significantly extend previous work by showing that the individual finger movements of one hand can be successfully distinguished with ECOG signals. In contrast to other multiclass BCIs that use well-separated body parts (e.g., hand/tongue/foot) or other forms of imagery (e.g, mental arithmetic, rotation) [6], [7], [8], we focus on a specialist area, i.e., the hand and fingers. Thus our results, in addition to improving the available bandwidth, may also aid in the development of intuitive prostheses that replace hand function.

From the perspective of understanding ECOG signals, our results show that high-frequency band-power features are the most informative for motor activity, and are sufficient for good classification performance even for a 5-class problem. From a machine learning perspective, we provide further evidence that sparse classification methods are robust against noise and data scarcity, and perform automatic feature selection.

Our study focuses on real movement of fingers. We believe this is an essential first step to characterizing changes in ECOG signals associated with motor activity, and may help form the basis of a “true” BCI, i.e., one which does not use explicit movement.

II. MEASURING AND REPRESENTING ECOG SIGNALS

A. Subjects

Electrocorticography (ECOG) is an electrical brain signal measurement technique used in patients suffering from intractable epilepsy. The procedure involes implantation of intracranial electrode arrays over the surface of the brain in order to localize seizure foci prior to surgical resection of the epileptic focus. In contrast to EEG, ECOG measurements show significantly higher spatial resolution and are unaffected by muscle activity and other artifacts.

Our subject pool is drawn from volunteering patients at the Harborview Medical Center, Seattle. Each patient had an implanted 8x8 platinum electrode array (Ad-Tech, Racine, WI), with 1 cm inter-electrode distance. The electrodes were embedded in silastic with 2.3mm diameter exposed (of a 4mm diameter electrode). Only patients with some peri-Rolandic coverage were included. Patients were studied 4-6 days after electrode placement to allow for recovery from the original surgery.

B. Experimental Protocol

We asked 6 subjects to move fingers of the hand contralateral to the grid placement, in response to visual cues on a computer screen. Subjects performed repeated movements of each individual finger of one hand for 2s-long intervals, interspersed with each other and with rest periods. Each finger was moved for 30 intervals, yielding a total of 150 data points spread over 5 classes of movement. ECOG data from a 64-electrode grid was recorded at 1000Hz and annotated with the stimuli presented to the user. The multi-purpose BCI2000 software [11] was used for presenting stimuli and recording data.

C. Spectral Features Encoding Movement

We used band power features calculated across each movement interval as our representation for each channel. Our previous work [9], [10] shows that 11-40Hz and 71-100Hz bands are very effective for localizing motor cortical areas corresponding to body parts, and for classifying signals measured during these movements. In other work [12], we have also explored the use of *very high band* power features, in the range 101-150Hz. For our present study, we used all 3 band power features, and our data representation consists of a 192-dimensional vector.

III. CLASSIFICATION METHODS

Based on our previous work [10], we used two classification methods: the Support Vector Machine (SVM), and a sparse variant of the SVM called the Linear Programming Machine (LPM). The methods are both linear binary classifiers, i.e., they assume linear separability of two-classes of data and attempt to find a hyperplane separating the data points belonging to the different classes. Linear classifiers are easier to train and interpret than other nonlinear methods, especially in the case of limited training data. We first describe the binary classifiers in detail and subsequently describe how multiclass problems can be solved by combining binary classifiers.

A. Classifiers

For data points $\mathbf{x_k}$ with labels $y_k \in \{+1, -1\}$, a linear classifier assigns the label $sign(\mathbf{w}^T\mathbf{x_k} + b)$, where the pair $(\mathbf{w}, b)$ are parameters of the classifier. Numerous optimization criteria can be used to estimate the parameters $(\mathbf{w}, b)$ from training data in a way that minimizes misclassified points. For the SVM classifier [13], the criterion is as follows:

$$min_{\mathbf{w},\xi,b} \quad \tfrac{1}{2}\|\mathbf{w}\|_2^2 + \tfrac{C}{K}\|\xi\|_1 \quad subject\ to$$

$$\begin{array}{rcll} y_k(\mathbf{w}^T\mathbf{x}_k + b) & \geq & 1 - \xi_k & and \\ \xi_k & \geq & 0 & for\, k = 1, ..., K \end{array} \qquad (1)$$

We use $\|\cdot\|_1$ to represent the l_1-norm (i.e. $\|\mathbf{w}\|_1 = \sum |w_i|$), and $\|\cdot\|_2$ to be the Euclidean or l_2-norm. The "slack variables" ξ_k have been introduced to trade off error from misclassified points with the quality of the classifier (the *margin* $\frac{1}{2}\|\mathbf{w}\|_2^2$, see [13] for more details). This tradeoff is mediated by the free parameter C, which is chosen empirically using cross-validation.

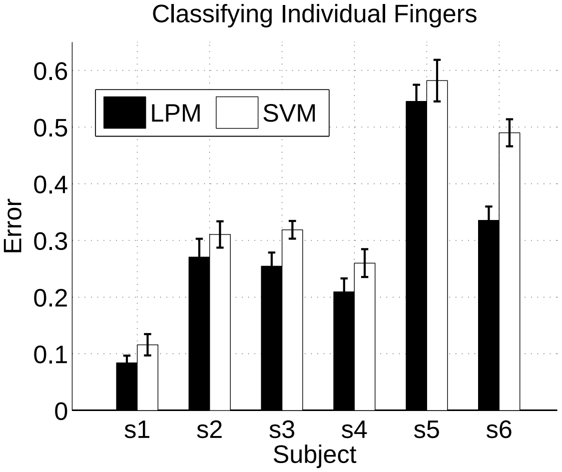

Fig. 1. **Classifying finger movement activity:** The figure shows the 5-class cross-validation error for the LPM and SVM classifiers, across 6 subjects. The results show that a high degree of accuracy is possible in distinguishing individual finger movements using ECOG. Also, the LPM consistently outperforms the SVM. (Chance level for a 5-class problem is 80% error.)

The quadratic program used for the SVM can be converted to a linear programming problem by replacing the l_2-norm on the regularizer with the l_1-norm [14]. This classifier is known as the Linear Programming Machine (LPM). The resulting linear program results in *sparse* weight vectors $\mathbf{w}$, i.e., vectors with most components zero or close to zero. This is a useful feature, since the nonzero components can be interpreted as the result of a *feature selection* step, thus making the classifier more robust to noise and overfitting.

The LPM is the solution to the following optimization problem:

$$min_{\mathbf{w},\xi,b} \quad \tfrac{1}{N}\|\mathbf{w}\|_1 + \tfrac{C}{K}\|\xi\|_1 \quad subject\ to$$

$$\begin{array}{rcll} y_k(\mathbf{w}^T\mathbf{x}_k + b) & \geq & 1 - \xi_k & and \\ \xi_k & \geq & 0 & for\, k = 1, ..., K \end{array} \qquad (2)$$

The free parameter C now controls the tradeoff between sparsity of the weight vector and the errors made by the classifier–a high value of C would impose a more severe penalty on misclassifications, and favor sparseness of the weight vector $\mathbf{w}$.

B. Multi-Class Classification and Error Measures

Binary classifiers can be combined to solve multiclass problems using one of two popular schemes. These are:

- One versus All (OVA): A separate classifier is trained to distinguish data points of each class from data points of all other classes. On a test data sample, the classifier with maximum positive output "wins", i.e., the data point is labeled with that class.
- All versus All (AVA): A separate classifier is trained for every pair of classes. For a test point, each classifier gets one vote with regard to the data point's class labels,

and the class with the maximum number of votes is the label assigned to the test sample.

Recent research [15] indicates that for strong well-tuned binary classifiers, OVA and AVA are comparable in performance, with neither method offering a clear advantage. We implemented both methods in our experiments, but found that OVA suffered from unbalanced sample sizes, a problem exacerbated by the extreme scarcity of data (only 30 samples of over 100 dimensions per class). We therefore report results only with AVA multiclass implementations.

C. Model Selection and Evaluation

We use 5-fold crossvalidation error as our performance measure. A nested cross-validation routine is used to estimate the free parameter C for our classifiers, i.e., for each outer crossvalidation fold, a 5-fold crossvalidation error on the training data is minimized to select C. The trained classifier is then evaluated on the unseen test fold. The entire nested crossvalidation routine is repeated 10 times and the average error over all runs is presented as a measure of classifier performance.

We implemented the LPM with the use of matlab's *linprog* linear optimization package. For SVM, we used the LIBSVM package [16].

IV. FINGER CLASSIFICATION RESULTS

A. 5-Class Classification error

Figure 1 shows the 5-class cross-validation error for 6 subjects, alongwith the standard deviation across repeated measurements. We see that there is a substantial amount of discriminative information in ECOG data regarding finger movements. The average error across 6 subjects is 23% for the LPM (for reference, a random classification strategy would have an error of 80%). Also, the LPM classifier consistently outperforms the SVM classifier. Subject 5 had poor electrode coverage of the cortical area for hand representation, possibly explaining the high classification error.

Other results (not shown) indicated that both classifiers performed significantly better when using the all-versus-all method for combining binary classifiers.

Figure 2 shows a closer picture of the classifier confusion (i.e., the fraction of data points labeled $finger_1$ that were classified as $finger_2$). The rows and columns are, in order, the thumb, forefinger, middle finger, ring finger and little finger. From the confusion matrices it is clear that ECOG signals during thumb movement are easily distinguishable from signals during movement of other fingers. In addition, the 2nd and 3rd, as well as the 4th and 5th fingers show some overlap in activity and label prediction confusion.

This is in keeping with both the classical anatomical representation of the hand, and studies investigating behavioral correlation and interdependence of finger movements [17]. An interesting open question is whether the behavioral correlations have cortical origins, thus explaining away some of the classification confusion as well.

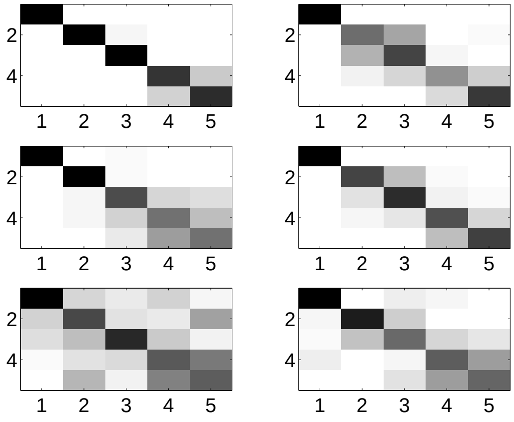

Fig. 2. **Classifier Confusion Matrices:** Shown are the confusion matrices for the classifier on each of the 6 subjects. The numbers indicate { thumb, fore, mid, ring, little} fingers, in order. We note that there are similarities to known behavioral confusion (e.g., behaviorally, the thumb is known to be independent of the other fingers).

B. Feature Selection

We explored two feature selection schemes: in one, we restricted the feature encoding to bandpower in only one of the three bands, in order to contrast the discriminative power of the individual bands against that of the combination. Figure 3 shows the result of this experiment, and indicates that with the high-band or very-high-band power alone, we can effectively predict finger movement as well as with the combination, indicating that the other two frequency bands are redundant. Further, it is clear that the low-frequency band power, although informative, is not quite as good a predictor as the other two bands. This is also in keeping with previous work on localizing hand representation [9], where we saw that the low-frequency band power changes were more diffuse, while high-frequency power changes were more focal in nature. Since the hand representation in the cortex is small compared to our electrode density, and individual finger representations are known to overlap, it is not surprising that the low-frequency bandpower feature performed poorly.

The second feature selection strategy we tried was the *active feature selection* strategy used by Mitchell et al. [18], who report significant performance improvement over conventional feature selection. The idea of feature selection is to rank features according to their discriminative power and select the top k% (or the ones that cross a particular threshold). However, when training data is scarce, this scheme may overfit and select some noisy features. To alleviate this, the authors compared each class to baseline data, and selected those features that were most active with reference to baseline.

In our case, the use of the sparse classifier eliminates the need for an additional feature selection step. This is clearly seen in contrast with the non-sparse SVM method in Figure 3-(b), which shows average error versus the fraction of features used. The LPM does not benefit significantly from a reduction in the number of noisy features, but the

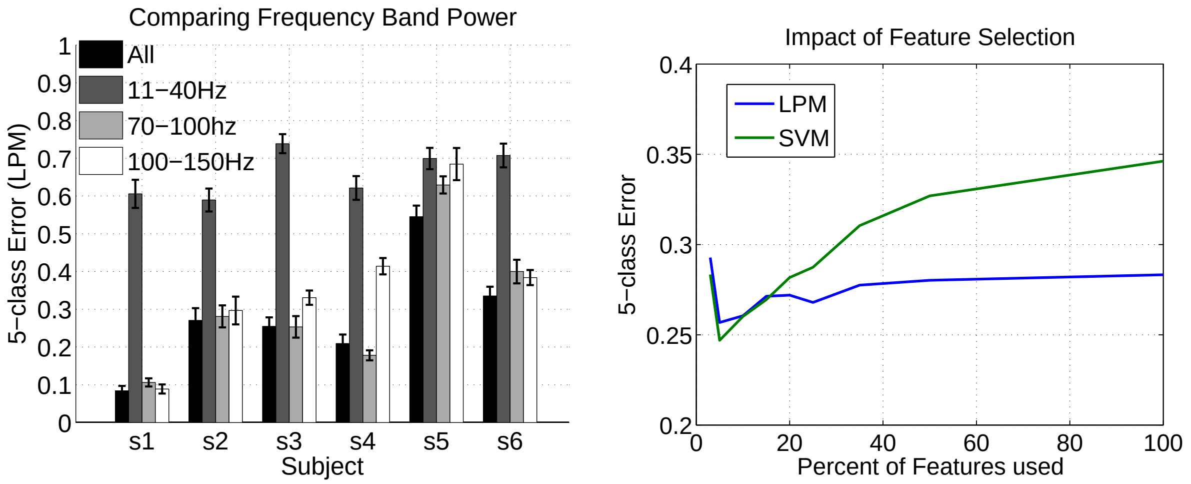

Fig. 3. **Impact of spectral and spatial feature selection on classification error:** (a) The figure on the left compares 5-class error for each subject when each frequency band is used in isolation, and when all features are combined. The two higher frequency bands individually contain all the information available in the combined representation. (b) The figure on the right shows the average error across all subjects when a subset of channels are used, chosen by their difference from baseline activity. The sparse classification method (LPM) performs well even without feature selection.

SVM classifier error drops significantly. These results further support the findings of our previous work [10], where we found that the sparse methods automatically selected as few as 20% of the features in a hand-tongue classification task.

V. CONCLUSION

We have shown that classifying individual fingers is possible with high accuracy; a result that is very useful for the development of an expressive high-bandwidth BCI system. The high classifiability of a 5-class problem involving cortical areas that are very close to each other indicates that even more information could potentially be extracted from ECOG signals. Our experience supports the use of high-band power features and sparse classifiers, combined in an all-pairs fashion, for multiclass classification of ECOG signals. As part of future work, we intend to address the following closely related questions:

- *Tracking finger movements through time:* Can we use the learned classifiers on shorter windows of time, in order to track the instantaneous positions of the fingers? Do high-frequency changes reflect instantaneous position or movement, and can one be inferred from the other?
- *Superposition:* Will the classifiers correctly predict activity during a concerted effort between fingers, e.g., freeform movement, or a grasp/pinch motion?
- *Imagined Movement:* How does imagined movement compare to real movement in terms of classifiability?

REFERENCES

[1] JR Wolpaw and DJ McFarland. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proc Natl Acad Sci U S A*, 101(51):17849–54, 2004.

[2] B Blankertz, G Dornhege, M Krauledat, K-R Mueller, V Kunzmann, F Losch, and G Curio. The Berlin Brain-Computer Interface: EEG-based communication without subject training. *IEEE Trans. Neural Sys. Rehab. Eng*, 14(2), 2006.

[3] EC Leuthardt, KJ Miller, G Schalk, RPN Rao, and JG Ojemann. Electrocorticography-based brain computer interface-the seattle experience. *IEEE Trans Neur Sys Rehab Engg*, 14 (2), 2006.

[4] LR Hochberg, MD Serruya, GM Friehs, JA Mukand, M Saleh, AH Caplan, A Branner, D Chen, RD Penn, and JP Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, 442, 2006.

[5] J. R. Wolpaw et al. Brain-computer interface technology: a review of the first international meeting. *IEEE Trans Rehab Engg*, 8:164–173, 2000.

[6] Obermaier B, Neuper C, Guger C, and Pfurtscheller G. Information transfer rate in a five-classes brain-computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 9, 2000.

[7] G Dornhege, B Blankertz, G Curio, and K-R Mueller. Boosting bit rates in non-invasive eeg single-trial classifications by feature combination and multi-class paradigms. *IEE Trans Biomed Engg*, 51(6), 2004.

[8] A Schloegl, F Lee, H Bischof, and G Pfurtscheller. Characterization of four-class motor imagery eeg data for the bci-competition 2005. *J. Neur Eng*, 2(4), 2005.

[9] KJ Miller et al. Electrocorticographic spectral changes with motor movement. *J Neurosci*, 2007, to appear.

[10] P. Shenoy, K.J. Miller, J. Ojemann, and R. Rao. Generalizable features for electrocorticographic bcis. Technical Report 2006-11-03, UWCSE, 2006.

[11] G Schalk, DJ McFarland, T Hinterberger, N Birbaumer, and JR Wolpaw. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE TBME*, 2004.

[12] KJ Miller et al. Real-time mapping of hand motor areas using high-frequency spectral changes. *Submitted*, 2007.

[13] V.N. Vapnik. *The Nature of Statistical Learning Theory*. Springer Verlag, New York, 1995.

[14] K.R. Müller, Krauledat M., Dornhege G., Curio G., and Blankertz B. Machine learning techniques for brain-computer interfaces. *Biomed. Tech*, 49(1):11–22, 2004.

[15] R Rifkin and A Klautau. In defense of one-vs-all classification. *JMLR*, 5:101–141, 2004.

[16] Chih-Chung Chang and Chih-Jen Lin. *LIBSVM: a library for support vector machines*, 2001. Software available at url-http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[17] C Haeger-Ross and MH Schieber. Quantifying the independence of human finger movements: Comparisons of digits, hands and movement frequencies. *J Neurosci*, 20(22), 2000.

[18] Mitchell T., Hutchinson R., Niculescu S., Pereira F., Wang X., Just M., and Newman S. Learning to decode cognitive states from brain images. *Machine Learning Journal*, Vol. 57, Issue 1-2:145–175, 2004.