

## A Hybrid SVM/HMM based System for the State Detection of Individual Finger Movements from Multichannel ECoG Signals

Ibrahim Onaran, N. Firat Ince, A. Enis Cetin, Fellow, *IEEE*, Aviva Abosch

**Abstract**— A hybrid state detection algorithm is presented for the estimation of baseline and movement states which can be used to trigger a free paced neuroprosthetic. The hybrid model was constructed by fusing a multiclass Support Vector Machine (SVM) with a Hidden Markov Model (HMM), where the internal hidden state observation probabilities were represented by the discriminative output of the SVM. The proposed method was applied to the multichannel Electrocorticogram (ECoG) recordings of BCI competition IV to identify the baseline and movement states while subjects were executing individual finger movements. The results are compared to regular Gaussian Mixture Model (GMM)-based HMM with the same number of states as SVM-based HMM structure. Our results indicate that the proposed hybrid state estimation method out-performs the standard HMM-based solution in all subjects studied with higher latency. The average latency of the hybrid decoder was approximately 290ms.

### I. INTRODUCTION

Neuroprosthetics (NP) aim to restore communication and control capabilities of people with debilitating motor impairments. Several neuroprosthetic systems have been constructed to process invasively recorded neural signals such as single-unit neuronal activity (SUA) for the control of a cursor on a computer screen, or for the control of a robotic arm. In most of these systems, the decoding process was restricted to predefined time intervals in which the state of the subject was altered by external cues limiting the flexibility of the constructed system.

In order to build a system that serves a subject's free will, the state of the brain activity needs to be determined to avoid undesired movement and to obtain accurate results for controlling an external device. For this particular purpose, in a free-paced NP, the states that need to be estimated dynamically are generally, (1) baseline (idle), (2) planning and (3) movement execution. Several attempts have been made to decode the dynamic state of the subject from neural activity [1-4]. The estimation of baseline, movement planning, and movement execution states from SUA was initially studied in [1] while non-human primates were executing directional hand movements in response to an

externally cued paradigm. In this work by Krishna and colleagues, neuronal firing rate, computed in fixed-size windows, was used as an input to a Bayesian state estimator, with the firing rates associated with each direction and with each state modelled with a Poisson distribution. A maximum likelihood (ML) classifier was then stamped each time window and the classification outputs were streamed to a finite state machine (FSM) for estimating the state of the subject. The FSM operated on *ad hoc*-derived transition rules. This work was extended by Achtman et al. [2], who constructed a two-stage decoder that was also based on an FSM. In contrast to [1], a growing window size was used in [2] to estimate from the neural data both the state and the direction of target. These investigators reported a 350-ms average latency in detecting executed movement directions.

Kemere et. al. [3] used a Hidden Markov Model (HMM) coupled with a state-dependent Poisson firing model instead of an FSM. These investigators demonstrated that using the *a priori* likelihood of the HMM states to first detect the onset of movement planning and then to calculate the ML target, results in substantial increases in performance relative to the FSM. The average latency for this study was approximately 330 ms with an 84% success rate in detecting executed movement directions.

Recently, Huang and Andersen used local field potentials (LFPs) recorded from the parietal cortex of primates during a directional reaching task, for a state decoding application [4]. This study demonstrated the feasibility of detecting state transitions from the oscillatory neural activity (LFPs) recorded with penetrating microelectrodes.

Recent studies indicate that HMM-based solutions provide better results than FSM-based solutions that are based on *ad-hoc* decision rules. A common setup shared by these studies is the externally cued paradigms that were used to alter the state of the subject in a controlled manner. In this work, we aimed to decode the movement and resting states of individual fingers from multichannel ECoG recordings. This is different from the previous studies that have focused on movement of the *entire* hand. In our scheme, we developed a new hybrid decoding system based on the fusion of SVM and HMM structures. Our discriminative/generative approach accepted input features computed with common spatial patterns in different frequency bands and returned the likelihood of one of the states of interest. To the best of our knowledge, this is the first study that explores the detection of movement execution and resting states of individual finger

I.O. and A.E.C are in the Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey. (e-mail: onaran@ee.bilkent.edu.tr, cetin@bilkent.edu.tr).

N. F. I. is in the Electrical and Computer Engineering Department, University of Minnesota, Minneapolis, MN 55455 USA, (e-mail: firat@umn.edu).

A. A. is in the Department of Neurosurgery, University of Minnesota, Minneapolis, MN 55455 USA, (e-mail: aabosch@umn.edu).

movements from ECoG recordings. Another novelty of our study is that it explored the success of decoding sequential movements in a continuous fashion rather than movements in a trial-based paradigm.

We used the hybrid decoding system to identify baseline and movement states. Moreover, we compared it to the conventional HMM method that has enjoyed widespread application in the field. A schematic diagram describing our signal-processing framework is depicted in Fig. 1. Below we first describe the dataset and the experimental paradigm. Next, we explain our signal-processing framework in detail. Finally, we provide experimental results and discuss these.

## II. METHODS AND MATERIALS

### A. ECoG Data

We used multichannel ECoG data from BCI Competition IV, recorded during finger flexion. This data set was acquired from three epileptic patients at Harborview Hospital in Seattle, WA. The electrode grid was placed on the cortical surface. Each electrode array contained either 48 (8x6) or 64 (8x8) platinum electrodes, each of 4-mm diameter. Electrode contacts were embedded in a silicon mat, and were spaced 1-cm apart. Synamps2 amplifiers (Neuroscan, El Paso, TX) were used to digitize and amplify the ECoG signal. The finger index to be moved was indicated with a cue on a computer monitor placed at the bedside. Each cue lasted two seconds and was followed by a two-second rest period, during which the screen was blank. Subjects moved one of five fingers 3-5 times during a cue period, for a total of 10 minutes for each subject [5]. The movements were continuous not trial based. Only the position of the fingers was available to us, and was used to distinguish between baseline (resting) and movement states. Consequently, this posed a great challenge in detection of these arbitrary movement executions as no information about the cue and go signal was available to us for our analysis. An exploratory analysis established that the duration and interval between consecutive finger movements varied dramatically. We used for analysis those segments in which each movement lasted a minimum of 1000ms and consecutive movements were separated by at least 800ms.

### B. Common Spatial Patterns

As in any learning process, the generalization capacity of a model decreases with the increasing dimensionality of the input data. Moreover, the complexity and execution time of decoding algorithms increases with the number of channels of input data. Therefore, a dimension reduction algorithm must be employed to decrease the dimensionality of data. We applied a Common Spatial Patterns (CSP) [6] algorithm on band-pass filtered multichannel ECoG signals in order to reduce these into a few virtual channels. Specifically, ECoG data from each subject was filtered in 1-4, 7-13, 16-30, and 65-200Hz frequency bands. Next, each band was

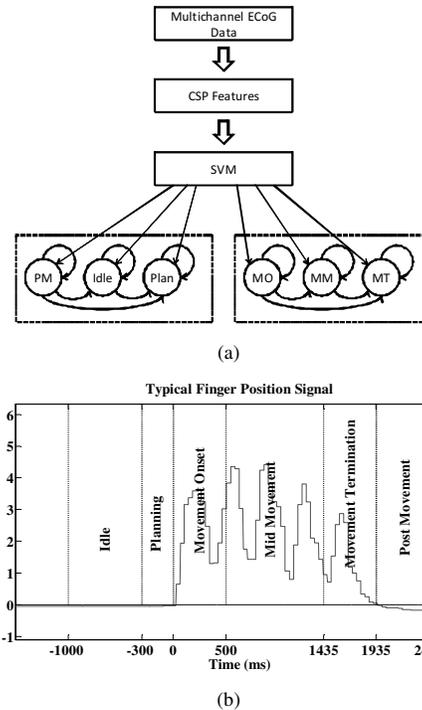


Fig. 1. (a) Multichannel filtered (1-4, 7-13, 16-30, and 65-200 Hz) ECoG data is fed into a CSP algorithm to reduce channel size. Each band is reduced into 4 virtual channels. Using the 16-dimensional CSP features, a multiclass SVM classifier is trained to distinguish between resting, planning, movement onset, mid-movement, movement termination and post-movement segments given in (b). These segmentations were derived by aligning the data to movement onset and movement termination. The SVM output probabilities were fed into two HMM models as observation probabilities of the hidden states. Prior and transition probabilities were computed from the training sequence using forward-backward method, where the model is restricted to left-to-right transitions only.

transformed into four virtual channels by the CSP algorithm, by taking the first and last two eigenvectors. We computed the spatial projection using

$$\mathbf{X}_{\text{CSP}}[n] = \mathbf{W}^T \mathbf{X}[n] \quad (1)$$

where the columns of  $\mathbf{W}$  are the eigenvectors representing each spatial projection and  $\mathbf{X}[n]$  is the multichannel ECoG data. The eigenvectors of the CSP algorithm were estimated via generalized eigenvalue decomposition by contrasting the covariance matrices of the resting and movement segments of the training data. Consequently, the CSP output maximized or minimized the variances of the resting and movement regions in the estimated virtual channels. The variance of each channel was computed in 250-ms windows moving with a 50-ms time step. Finally, the variances were log transformed and concatenated across all four frequency bands forming a 16-dimensional feature vector for each time shift.

### C. Hybrid HMM-SVM Structure

In order to estimate resting and movement states from the recorded neural data, we built a hybrid discriminative/generative decoder based on the fusion of

HMM with SVM. HMMs are widely used in speech processing and have been successfully applied to dynamic state decoding of neural data. Detailed descriptions of this method and its applications have been published elsewhere [7]. Because it is a generative method, the HMM structure lacks discrimination capability: each model is trained independently from the other competing models. Moreover, observation probabilities are generally modeled by Gaussian Mixture models (GMM), which fail to represent the distribution of the features in high dimensional space in the presence of a low amount of training data and/or outliers. We therefore aimed to replace the observation probabilities of internal states of the HMM with the posterior probability output of a multiclass SVM. Specifically, rather than using a GMM, the extracted features were fed to a multiclass SVM that was tuned to separate the distribution of the internal states. However, such an approach requires the labels of the features belonging to each state so that the SVM classifier can be trained. In this scheme, we constructed six different states by aligning the neural data with respect to movement onset and termination. These states consisted of the following six periods: i) resting, ii) movement planning, iii) movement onset, iv) mid-movement, v) movement termination stage, and vi) post-movement stage. A schematic diagram representing these alignments and their duration is given in Fig. 1.b. Because there was no exact timing information for the planning period, we used the 400-ms window preceding each movement onset as the planning state (P). The 400-ms segment immediately following each movement termination was defined as the post-movement state (PM). The interval between PM and P was defined as the resting segment. Movement was segmented into three different states, with the first 400ms of each movement defined as movement onset (MO). The 400-ms segment immediately preceding cessation of movement was defined as the movement termination state (MT). The interval between MO and MT was defined as the mid-movement state. We labeled the features originating from each state in the continuous training data and then fed them into the multiclass SVM for discrimination. Since the duration of the resting and mid-movement states was variable, the number of feature vectors that we extracted from these segments was much higher than for the other states, causing a bias in the decision boundary of the SVM classifier. Consequently, we reduced the number of samples for resting and mid-movement states in order to compensate for the variability in numbers of samples for each state. Specifically, the majority class was down-sampled by randomly eliminating its samples. The SVM module provided an estimated posterior probability for each state by using a *one against the other* classification strategy. A radial basis function was used as the kernel of the SVM. The output of the SVM module was then used in conjunction with the Forward-Backward algorithm to estimate the transition probabilities of the HMM. We used the LibSVM toolbox to implement the

	Hybrid Decoder	HMM
Subject 1	91.5	89.2
Subject 2	89.2	88
Subject 3	92.7	91.6

Table 1. The state decoding accuracies of the hybrid and traditional HMM based methods with 60 training trials using a decoding sequence length 10.

multiclass SVM [8] and the HMM toolbox of [9] to build the hybrid decoder. It should be noted that this procedure differs from the traditional HMM training, in which the observation and transition probabilities are altered in each iteration of the standard Expectation-Maximization (EM) algorithm. In our case, the observation probabilities were the SVM outputs, and these were fixed during the iterative estimation of transition probabilities. The HMM model had three hidden states. In each state, the observation probabilities were represented with three mixtures. Only left-to-right transitions were allowed in both hybrid and HMM, as depicted in Fig 1.

We tested our hybrid decoding system and the traditional HMM algorithm on the ECoG data derived from the three subjects of BCI Competition IV, described in Section II. In contrast to those studies that have decoded transition from baseline to planning/movement, our challenge involved decoding transitions from movement to a resting/baseline state, as well. In order to decode the dynamic state of a subject, a sequence of observations is needed. Unlike trial-based experiments, the data we used contained no prescribed start and end points. In such a situation, a fixed segment of the data, which is shifted along the signal, is generally used to execute the state decoders. The use of long data segments can cause large latencies and numerical overflow of the output. Consequently, we studied the effect of different sequence lengths, for example, 5, 10, 15 and 20, on the estimation of the resting and movement states. After decoding each sequence with the constructed models, the model with the maximum posterior probability was used to determine the class of the feature sequence. Moreover, we executed several experiments with various training-set sizes, in order to examine the robustness of each algorithm against the limited amount of training data. We trained the algorithms using 10 to 70 train trials by increasing the set size by ten.

### III. RESULTS

The average classification accuracies of the hybrid and HMM methods are listed in Table 1. We observed that for all subjects studied, the hybrid SVM-HMM decoder provided better decoding accuracies than the traditional HMM method. On average, the detection accuracy of the hybrid method was 91.2%, whereas the HMM solution provided 89.6% decoding accuracy.

The average decoding accuracies of each method with a varying number of training trials is given in Fig. 2 (a). We observed that the hybrid decoder provided superior decoding accuracies with a low number of training trials, and its

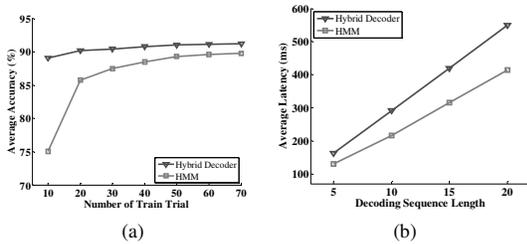


Fig. 2 Average accuracy vs. the number of train trial with decoding length 10 (a). Average latency vs. decoding sequence length (b).

performance slowly increased with increasing the training set size. In contrast, the accuracy of HMM was quite poor when using a low number of training trials. In contrast to the hybrid decoder, the accuracy of HMM rapidly improved with increasing training-set size, ultimately stabilizing after 50 training trials.

We studied decoding accuracy as a function of decoding sequence length. We observed that the decoding results were quite poor with a sequence length of five and improved rapidly by increasing the sequence length to ten. The maximum decoding results were obtained with sequence lengths of 10 and 15 in both methods, which corresponded to time windows of approximately 700 and 950ms, respectively. The average latency of each method versus the decoding sequence length is given in Fig. 3 (b). We observed that the latency of HMM was superior to the hybrid decoder. For a sequence length of 10, the latency for the hybrid and HMM were 290 and 215 ms, respectively. Although slightly better results were obtained using a sequence length of 15 with the hybrid decoder, we observed that the latency increased dramatically from 290 to 410ms.

The temporal decoding accuracies for a representative subject at movement onset and termination are shown in Fig.3. We observed that the decoding results at movement onset had a sharp transition compared to movement termination. We also noted that the decoding errors and latencies were higher at movement termination, as compared to movement initiation. These observations indicate that decoding state transitions from movement to resting state poses new challenges. In the subjects we studied, movement onset was associated with a burst of gamma spectrum activity, which slowly decreased towards the end of the movement. There was no similar pattern observed at movement termination. This could in part explain the lower accuracy and the larger latency that characterized movement termination.

#### IV. CONCLUSION

We report here a hybrid decoder based on the fusion of SVM and HMM for dynamic state detection based on data derived from multichannel ECoG recordings during consecutive movements of individual fingers. We have demonstrated experimentally that the latency of state

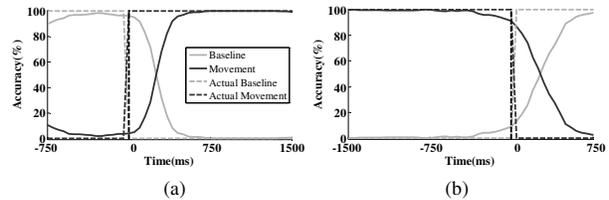


Fig.3 Average accuracy vs. time for subject 3 aligned to the movement onset in (a), and to the movement termination in (b). A hybrid decoder with a decoding sequence length of 15 was used.

decoding using ECoG data during finger movements is comparable to that obtained using SUA data during directional hand movements. We compared our method to the traditional HMM technique. The hybrid decoder outperformed the HMM technique in all three subjects studied. The main advantage of using SVM within the hybrid decoder is that the posterior probability of each state is estimated simultaneously and tuned for discrimination. This advantage might overcome the lack of discriminative capability of HMMs, as each model is trained independently from the other competing models. Moreover, the higher generalization capacity of SVM due to the large margin makes the algorithm a good candidate for applications in which a limited number of training trials exists on which to base estimates of the model parameters. However, such an approach requires supervised training in order to estimate the state discriminators, which is automatically accomplished by the traditional HMM.

#### Acknowledgements

This study was supported in part by a grant from the University of Minnesota Interdisciplinary Informatics Program (UMII) and the National Scientific Research Council of Turkey (TUBITAK).

#### REFERENCES

- [1] Krishna V. Shenoy et al., "Neural prosthetic control signals from plan activity," *NeuroReport*, vol. 14, no. 4, pp. 591-596, 2003.
- [2] Neil Ahtman et al., "Free-paced high-performance brain-computer interfaces," *Journal of Neural Engineering*, vol. 4, no. 3, pp. 336-347, September 2007.
- [3] Caleb Kemere et al., "Detecting Neural-State Transitions Using Hidden Markov Models for Motor Cortical Prostheses," *J Neurophysiol*, no. 100, pp. 2442-2452, June 2008.
- [4] Eun Jung Hwang and Richard A. Andersen, "Brain-control of movement execution onset using LFPs in posterior parietal cortex," *J Neurosci.*, vol. 29, no. 45, pp. 14363-14370, November 2009.
- [5] Kai J. Miller and Gerwin Schalk. (2008, June) Prediction of Finger Flexion 4th Brain-Computer Interface Data Competition. [Online]. [http://www.bbci.de/competition/iv/desc\\_4.pdf](http://www.bbci.de/competition/iv/desc_4.pdf)
- [6] Koles ZJ, Lazar MS and Zhou SZ (1990) Spatial patterns underlying population differences in the background EEG. *Brain Topogr* 2: 275-284.
- [7] L.R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257-286, February 1989.
- [8] Chih-Chung Chang and Chih-Jen Lin. (2001) A library for support vector machines. <http://www.csie.ntu.edu.tw/~cjlin/lib>
- [9] Kevin Murphy. (1998) Hidden Markov Model (HMM) Toolbox for Matlab. <http://www.cs.ubc.ca/~murphyk/Software/HMM/hmm.html>