Modeling the Relation from Motor Cortical Neuronal Firing to Hand Movements Using Competitive Linear Filters and a MLP

Sung-Phil Kim[1], Justin C. Sanchez[1], Deniz Erdogmus[1]
Yadunandana N. Rao[1], Jose C. Principe[1], Miguel Nicolelis[2]
[1] Electrical Engineering Department, University of Florida, Gainesville, FL 32611, USA
[2] Department of Neurobiology, Duke University, Durham, NC 27710, USA

Abstract – **Recent research has demonstrated that linear models are able to estimate hand positions using populations of action potentials collected in the pre-motor and motor cortical areas of a primate's brain. One of the applications of this result is to restore movement in patients suffering from paralysis. To implement this technology in real-time, reliable and accurate signal processing models that produce sufficiently small error in the estimated hand positions are required. In this paper, we propose the hybrid model approach that combines competitive linear filters with a neural network. The mapping performance of our approach is compared with a single Wiener filter during reaching movements. Our approach demonstrates more accurate estimations.**

I. INTRODUCTION

Brain-machine interfaces are a developing technology that aims to transfer the intent of an individual to a machine. Our goal is to substitute physical control by electrical signals originating from the operator's brain with a variety of devices. These devices serve patients suffering from neurological disorders.

Nicolelis and colleagues [1] showed that linear and nonlinear time-delayed neural network (TDNN) models can predict the hand positions of a primate using the firing patterns of populations of cortical neurons. Large arrays of microelectrodes implanted in the pre-motor and motor cortical areas of a primate record the activity of populations of neurons. Spike-detection and sorting algorithms are used to process these analog potentials to determine the firings of single neurons. A count of the number of spikes within 100-msec windows is computed and fed into either the linear filter trained with least squares or the TDNN trained with conjugate gradient to match the x, y and z coordinates of primate's hand.

Other groups have also demonstrated neural control of devices using linear and nonlinear methods. Primate spiral tracing prediction by the population vector algorithm has been proposed by Moran and Schwartz [2]. Chapin and colleagues demonstrated prediction of lever pressing from ensembles of rat cortical neurons using a recurrent neural network (RNN) [3]. Neural cursor control using linear filters trained with the least squares has also been proposed by Serruya *et al.* [4].

However, it remains unknown which linear or nonlinear model produces a closer approximation of the target function that generates hand positions from cortical firing patterns. We also consider the feasibility of real-time implementation to produce the brain-machine interface. It is obvious that the linear model usually provides the best computational cost for hardware implementation. Yet, it often fails to find a more complex input-output mapping that captures details in output trajectories. Recently, our group has demonstrated that the nonlinear model is able to estimate hand position more accurately than the linear model during reaching movements [5].

In this paper, we propose a new approach to model the input-output mapping using a two-layer neural network whose inputs are fed from a bank of linear filters (see Fig. 1). This hybrid architecture shows superior performance compared to a single linear filter, and very close performance to the RNN in [5].

II. MODEL

Preliminary results from the input-output mapping of neuronal spike counts to hand positions have shown that it is reasonable to assume a nonlinear relationship between input and output. Therefore, a single FIR filter can be sub-optimal for this nonlinear system modeling.

Our modeling method is a "divide and conquer" approach. A complex nonlinear modeling task can be elucidated by dividing it into simpler linear modeling tasks and combining them properly. Fancourt *et al.* have proposed the multiple linear models to segment a nonstationary signal [6]. Assuming that a nonstationary signal is a combination of piecewise stationary signals, subsystems can adapt to separate stationary portions so that each linear filter specializes in a specific temporal segment. From the observations of characteristics of hand movement trajectories (see top figure of Fig. 2), we assume that the hand trajectory and the neuronal spike count patterns belong to two movement regimes, moving and stationary ones. In this case, a temporal segmentation using multiple linear models may be able to provide a better overall input-output mapping. If we switch on-line between models with an appropriate method, the

This work was supported by a seed grant from the College of Engineering, University of Florida and the DARPA project # N66001-02-C-8022.

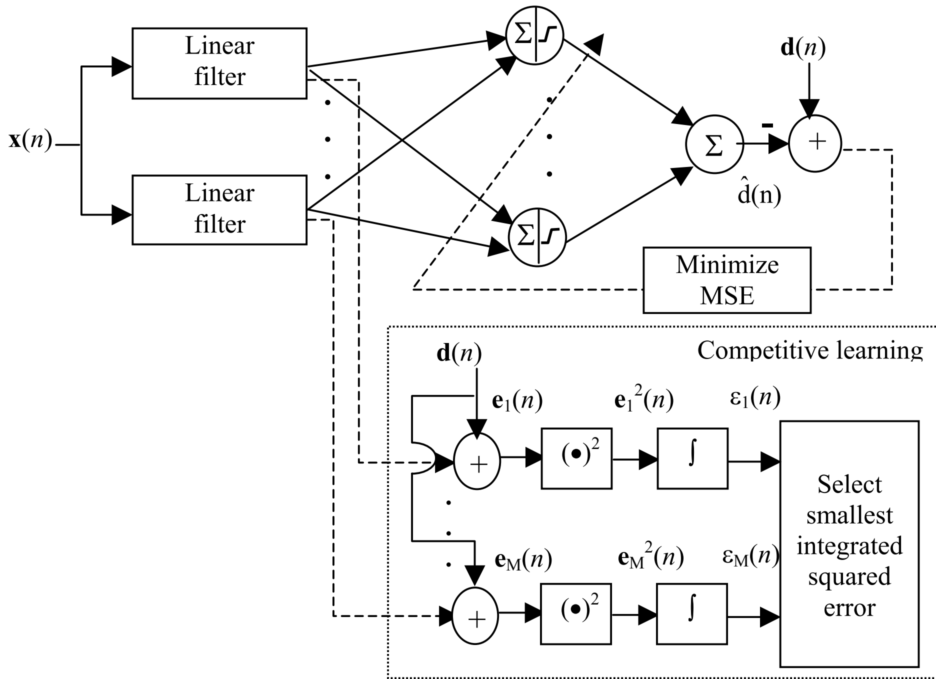


Figure 1. An overall architecture of the proposed model. The box illustrates the selection of the winner using integrated squared errors from each linear filter. Outputs from M trained linear filters are fed to a MLP, which is trained using the conjugate gradient algorithm. $d(n)$ denotes the desired response.

multiple models can estimate the desired response with sufficiently small error variance.

In our model, multiple adaptive finite impulse response (FIR) filters are built and trained by the normalized least mean square (NLMS) [7] learning rule to find mappings from the spike trains to hand movements. When an input is present, each filter produces an error between its output and the desired response. Each squared error is fed to an integrator to estimate recursively the expected value of the instantaneous squared error [6]. Output of the integrator for the i^{th} filter is,

$$\varepsilon_i(n) = (1-\mu)\varepsilon_i(n-1) + \mu e_i^2(n),\ i=1,\ldots,M \tag{1}$$

where, M is the number of filters, and μ is the feedback parameter that controls the integration time. The filter with the smallest integrated error finally wins competition. We can use a hard or a soft competition rule to update the weights of the winning filter. A hard competition rule updates only the winner at each time instance as,

$$\mathbf{w}_{winner}(n+1) = \mathbf{w}_{winner}(n) + \frac{\eta e_{winner}(n)\mathbf{x}(n)}{\gamma + \|\mathbf{x}(n)\|^2} \tag{2}$$

where, $\mathbf{w}_{winner}$ is the winning filter's weight vector, $\mathbf{x}(n)$ is the present input, $e_{winner}(n)$ is the error produced from the winning filter, η is the learning rate and γ is the small positive constant. A soft competition rule, on the other hand, updates all filters using a kernel as,

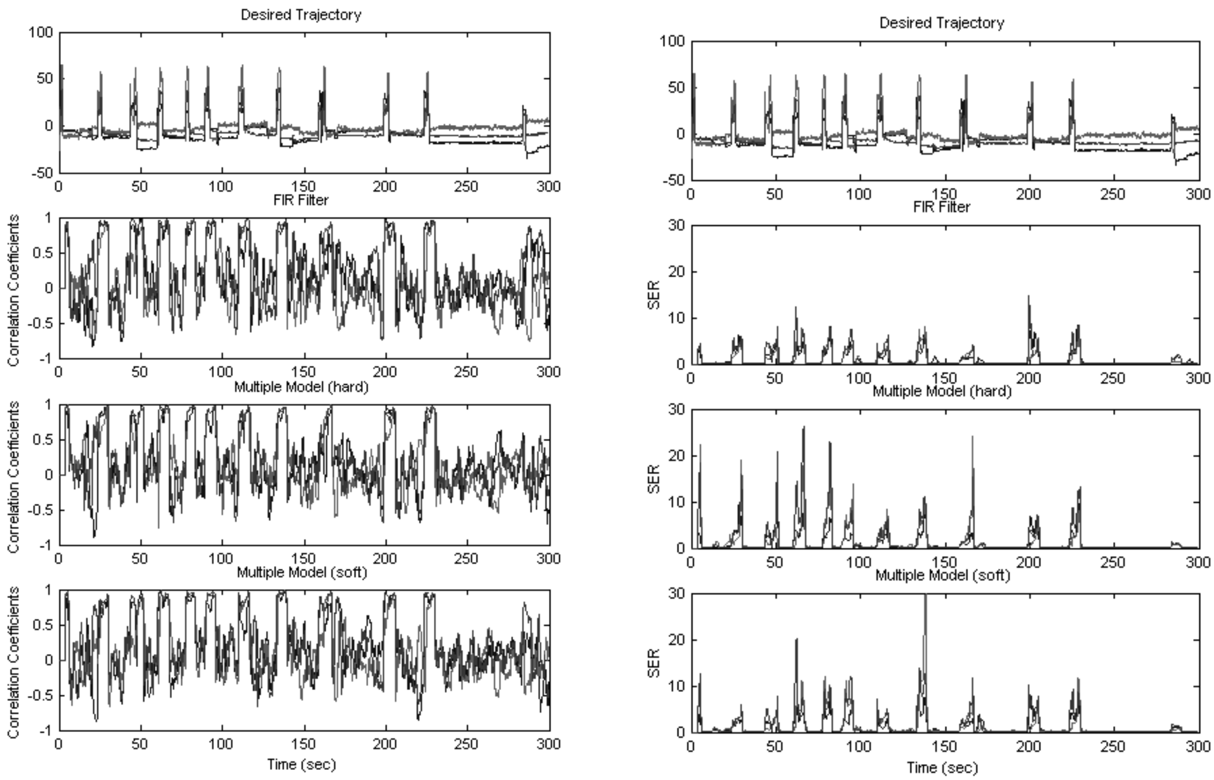
$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \frac{\eta(n)\Lambda_{i,j}(n)e_i(n)\mathbf{x}(n)}{\gamma + \|\mathbf{x}(n)\|^2},\ \text{i}=1,\ldots,M \tag{3}$$

where, i is the index of filter, j is the index of winner, and $\Lambda_{i,j}(n)$ is the kernel function. We use a Gaussian kernel as,

$$\Lambda_{i,j}(n) = \exp\left(-\frac{d_{i,j}^2}{2\sigma^2(n)}\right) \tag{4}$$

where, $d_{i,j}$ is the Euclidean distance between index i and j, which is equal to $|j-i|$, and $\sigma^2(n)$ is the kernel width, which decreases exponentially as n increases. Note that the learning rate also decreases with time. Fig. 1 depicts the selection of model using the integrated error.

Although multiple linear models are able to produce very accurate estimated hand positions during training, they always need a desired response to select the winner, because the error (including the desired response) is used as a selection metric. Since we are not able to access the actual hand positions in testing, the estimation of hand positions using only inputs and trained models must be considered. Previous work on segmentation [6] did not consider this case since the filters were being trained as predictors. Our approach feeds the outputs of the multiple linear models into a multilayer perceptron (MLP) to estimate hand positions during testing (without knowledge of the desired response after training). After multiple filters are trained, the same training samples are fed to the bank of filter with fixed weights. The outputs form input vectors of same size as the number of multiple filters to the MLP. The MLP is trained with the same desired response by the conjugate gradient algorithm effectively learning how to compute the desired hand positions from the filter outputs. This overall architecture is very similar to a focused TDNN, however it is



(a) (b)

Figure 2. The actual hand trajectory for x, y, and z coordinates are presented in each top plot for (a) and (b). The trajectories of the correlation coefficients (a), and the SERs (b) for 4-sec time window are depicted in the second, third and fourth subplots, for a single FIR filter, multiple models with hard competition, and multiple models with soft competition, respectively.

trained differently. The linear filters are competitively trained first, and then only the nonlinear mapping is trained. Conventional training of the TDNN [8] with backpropagation is a difficult task due to the attenuation of the error by the hidden layer processing elements (PEs). Alternatively estimating the a posteriori probability of the models given the data as in the mixture of experts [9] was also not successful.

III. Simulations

In this section, we present experimental results obtained with the proposed model using both hard and soft competition rules and compare those with results of a single FIR filter. A single FIR filter consists of the same number of delays as each filter in our multiple models.

A. Data

Synchronous, multichannel neuronal spike trains were collected at Duke University using owl monkeys (Aotus trivirgatus). Microwire electrodes were implanted in cortical regions where motor associations are known [1]. The firing times of single neurons were recorded while the monkey performed a 3-D reaching task. The monkey hand position was also recorded (with a shared time clock) and digitized with a 200Hz-sampling rate. The neuronal firings were binned in non-overlapping windows of 100ms, which represents the local firing rate for single neurons. These spike counts were directly used as inputs to the model to estimate hand positions. The digitized hand position signal was downsampled to 10Hz to synchronize with spike counts in time. In order to take the reaction time into account, the spike trains were delayed by 0.23 seconds with respect to the hand position.

B. Experiment Results

The spike counts of each of the 104 neurons were used to train our model with 10 linear filters, 20 hidden PEs and 3 linear output PEs to predict the x, y, and z coordinates of the monkey's hand. Preliminary results in [5] have shown that the "optimal" number of delays was 20 in terms of minimizing the MSE of test set. Therefore, each FIR filter contained 20 delays, and 3 outputs so that its weight vector has 6240 elements. The feedback parameter in the integrator was 0.1, which effectively represents a memory depth of 10. If we increase the feedback parameter, the choice of a winning filter during competition would be affected more by present prediction errors. A large feedback parameter may cause a biased selection of a winner, since the selection becomes more sensitive to present predicting ability with it [6]. A feedback parameter of 0.1, therefore, is a reasonable choice because it enables to compare the predicting performance approximately over ten points (1 sec). The number of filters and the number of hidden PEs were chosen empirically by examining the predicting performance on the test set. We restrained the size of them to prevent from resulting a huge network.

A training set of 20,000 consecutive bins (2,000 secs) of data was utilized. The linear filters were trained first and their weights were fixed. Next, the MLP was trained with the outputs of multiple filters for 500 epochs. Since the number of filters was 10, and the number of outputs from each filter was 3, total 30-dimesnional inputs were fed to the MLP. Training of the MLP is repeated with 100 random initial conditions and the minimum mean square error (MSE) solution is accepted. In testing, all the model parameters were fixed and 3,000 consecutive bins of new neuronal data were fed to the model to predict hand positions. The testing results are evaluated in terms of the correlation coefficients and the signal to error ratio (SER) between actual and estimated hand trajectories during 4-sec time windows (because movements take approximately 4 secs). The SER is defined as the power of the desired signal divided by the power of the estimation error. Since a high correlation coefficient does not account for a bias in the trajectories, a measure of SER should be used together to evaluate the performance more meaningfully.

In Fig. 2(a), the x, y, and z coordinates of hand trajectories during test are shown in the top subplot. Also shown in subplots of Fig. 2(a) are the correlation coefficients between the actual and the estimated hand trajectories for the single FIR filter, the multiple models with hard competition and the multiple models with soft competition, respectively. The figures show that the correlation coefficients are large when hand is moving. The cumulative correlation coefficients for the entire test set averaged over all coordinates were 0.64 ± 0.39 (standard deviation) for the single FIR, 0.72 ± 0.38 for the multiple models with hard competition, and 0.77 ± 0.38 for the multiple models with soft competition. In Fig. 2(b), the same desired trajectories are shown in the top subplot. The SERs for all three models are shown also. The SERs are windowed in the same way as the correlation coefficients. The maximum SERs reached values of 7.51, 15.96, and 19.25 for the single FIR filter, the multiple models with hard competition, and multiple models with soft competition, respectively. Cumulative SERs, which is averaged over all coordinates and the entire test set were 0.87 for the single FIR filter, 1.22 for the multiple model with hard competition, and 1.17 for the multiple model with soft competition. It is obvious that the multiple models significantly improve

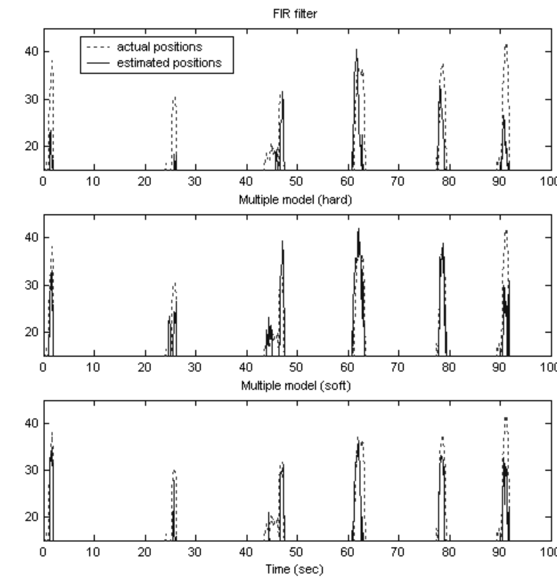


Figure 3. Peaks of hand trajectories for the z-coordinate.

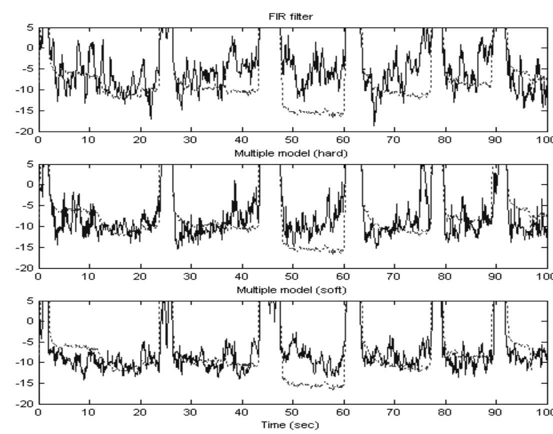


Figure 4. The actual and the estimated hand trajectories when hand is at rest position for the z-coordinate.

performance of estimation hand positions from neuronal spike counts compared to a single FIR filter. The issue of soft versus hard competition does not play a role in this data.

The peaks of the estimated hand trajectory superimposed on the actual trajectory are shown in Fig. 3. The first 1,000 samples (100 seconds) of hand trajectory were used to show how each model reaches the peaks of movements. Only three of six peaks are captured by the single FIR filter, while five peaks are captured by the multiple models. Fig. 4 shows the estimated hand trajectory superimposed on the actual trajectory when the monkey has the hand at rest. The same 1,000 samples as above were used. The multiple models with soft competition produced the hand trajectory with the least noise. This explains why the multiple models with soft competition provide the highest correlation coefficient.

The target accuracy of each model is further compared in Fig. 5 that shows the errors for three peak values (i.e., when the hand is reaching the target). In the figure, the target hand position is located at the origin. The distance between the estimated and the actual hand position associated with each direction (x, y, and z) is plotted on its respective axis. In all three plots, position is measured in terms of millimeters. We can see that the multiple models yields less deviation from the target hand position than the single FIR filter.

V. CONCLUSIONS

In this paper, we have developed hybrid architecture to predict the hand positions from neuronal spike populations. Our model combines competitive multiple models with a neural network framework. This approach provides significant improvements compared to the single linear model since it divides the eventually nonlinear mapping from neuronal data to hand positions into local linear mappings. Our model produced a close predicting performance to a RNN introduced in [5]. The cumulative correlation coefficient for the recurrent neural network was 0.75, the maximum SER reached a value of 34.19, and the cumulative SER was 1.46 according to experimental results on the exactly same test data used in this paper. The predicting performance of our model is comparable to the recurrent neural network except the maximum SER.

The soft competition rule results in more accurate estimations than the hard competition rule when the hand is at rest position, since every filter contains weighted information about all inputs and desired response. Although our model can reach the peaks of the trajectory with small error variance, it is unable to estimate the trajectory when hand is at rest position. This may be caused by the fact that spike train in the pre-motor and motor areas do not code for hand position at rest. However, since our model has a large number of weights, generalization may be an issue. Further study of the regularization methods can help us to compact a model fitted to our data.

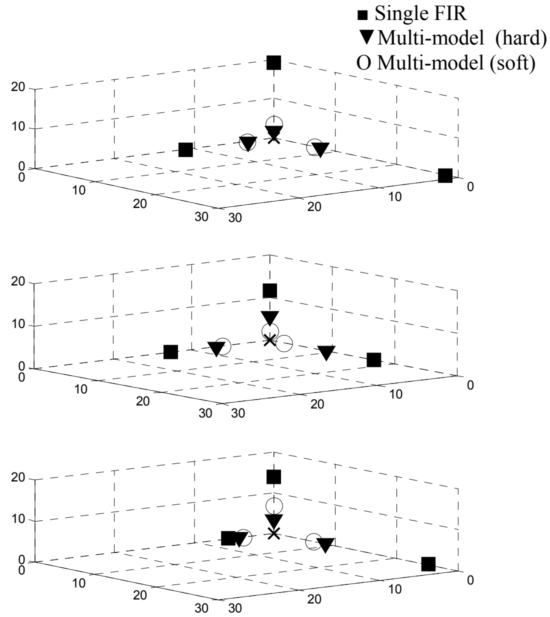


Figure 5. Estimation errors for three peak values. The absolute values of error (mm) in each direction are displayed on the respective axis.

REFERENCES

[1] J. Wessberg, R. Stambaugh, J.F. Kralik, P.D. Beck, M. Laubach, J.K. Chapin, J. Kim, J. Biggs, M.A. Srinivasan, M.A. Nicolelis, "Real-Time Prediction of Hand Trajectory by Ensembles of Cortical Neurons in Primates," Nature, vol. 408, no 6810, pp. 361-365, 2000.

[2] D.W. Moran, A.B. Schwartz, "Motor Cortical Activity During Drawing Movements: Population Representation During Spiral Tracing," Journal of Neurophysiology, vol. 82, no. 5, pp. 2693-2704, 1999.

[3] J.K. Chapin, K.A. Moxon, R.S. Markowitz, M.A. Nicolelis, "Real-Time Control of a Robot Arm Using Simultaneously Recorded Neurons in the Motor Cortex," Nature Neuroscience, vol. 2, no. 7, pp. 664-670, 1999.

[4] M.S. Serruya, N.G. Hatsopoulos, L. Paninski, M.R. Fellows, J.P. Donoghue, "Brain-Machine Interface: Instant Neural Control of a Movement Signal," Nature, vol. 416, pp. 141-142, 2002.

[5] J.C. Sanchez, S.P. Kim, D. Erdogmus, Y.N. Rao, J.C. Principe, J. Wessberg, M.A. Nicolelis, "Input-Output Mapping Performance of Linear and Nonlinear Models for Estimating Hand Trajectories From Cortical Neuronal Firing Patterns," Proceedings of NNSP'02, pp. 139-148, Martigny, Switzerland, 2002.

[6] C. Fancourt, J.C. Principe, "Temporal Self-Organization Through Competitive Prediction," Proceedings of ICASSP'96, pp. 3325-3329, 1996.

[7] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Upper Saddle River, NJ, 1996.

[8] S. Haykin, *Neural Networks: A Comprehensive Foundation*, New York: Toronto: McMillan: Maxwell McMillan Canada, 1994.

[9] A.S. Weigend, M. Mangeas, A.N. Srivastava, "Nonlinear Gated Experts for Time Series: Discovering Regimes and Avoiding Overfitting," International Journal of Neural Systems, vol. 6, pp. 373-399, 1995.