Decoding grasp aperture from motor-cortical population activity

Panagiotis K. Artemiadis, Gregory Shakhnarovich, Carlos Vargas-Irwin,
John P. Donoghue, and Michael J. Black

***Abstract*— The direct neural control of external prosthetic devices such as robot hands requires the accurate decoding of neural activity representing continuous movement. This requirement becomes formidable when multiple degrees of freedom (DoFs) are to be controlled as in the case of the fingers of a robotic hand. In this paper a methodology is proposed for estimating grasp aperture using the spiking activity of multiple neurons recorded with an electrode array implanted in the arm/hand area of primary motor cortex (M1). Grasp aperture provides a reasonable approximation to the hand configuration during grasping tasks, while it offers a large reduction in the number of DoFs that must be estimated. A family of state space models with hidden variables is used to decode each finger grasp aperture with respect to the thumb from a population of motor-cortical neurons. The firing rates of multiple neurons in M1 were found to be correlated with grasp aperture and were used as inputs to our decoding algorithm. The proposed decoding architecture was evaluated off-line by decoding pre-recorded neural activity from monkey motor cortex during a natural grasping task. We found that our model was able to accurately reconstruct finger grasp aperture from a small population of cells. This demonstrates the first decoding of continuous grasp aperture from M1 suggesting the feasibility for neural control of prosthetic robotic hands from neuronal population signals.**

I. Introduction

Direct neural control of external devices has recently become feasible. Recent results have demonstrated continuous neural control of devices such as computer cursors or simple robotic mechanisms using implanted electrodes in monkeys [1], [2], [3], [4] and humans [5]. These results are enabled by a variety of mathematical decoding methods that produce an estimate of the system "state" (e.g., hand position or velocity) from a sequence of measurements (e.g., the firing rates of a population of cells). In previous work, decoded movement signals were limited to the two or three-dimensional hand kinematics or grip force for simple robot grasping [4]. However hand posture, which will be essential for more complex dexterous tasks, has not previously been recovered. Here we show how the grasp aperture of each finger can be decoded resulting in a simple representation of the hand shape. Specifically it is shown that continuous parameters related to hand grasp aperture can be decoded from a small population of cells in primary motor cortex. This provides the proof of concept necessary for the prosthetic control of a robot hand in a neural interface system.

Analyzing a distributed neural control signal acting on a complex biological actuator, such as the hand, is a difficult technical challenge that has only recently become tenable. Two key technological developments were necessary: precise measurement of hand motion and simultaneous recording of the activity of a large number of cortical neurons. Motion capture technology addresses the first problem by noninvasively measuring the subtle kinematics of hand motion. Using high-speed cameras it is possible to reconstruct the position of reflective markers attached to the hand with sub-millimeter accuracy. Although motion capture has been successfully used in primate studies [6], relating this type of data to the activity of large populations of neurons, remains an extremely challenging issue that has not previously been addressed.

A large number of decoding algorithms have been proposed in the last decades. The population vector approach is the oldest method proposed in early 1980s [7]. Since then, a variety of methods have been proposed. Linear regression [8], artificial neural networks [3] and switching Kalman filter [9] are some of those. However, all the above methods were used for reconstructing hand kinematics (e.g. the position, velocity and acceleration of the hand), disregarding the individual finger movements during the experiments.

Recent work on the neurophysiology of hand movements supports the idea that the brain does not control each possible degree of freedom independently (e.g., each joint or muscle activation) [10]. Even when broken down to the simplest possible level (the activation of single neurons) the cortex appears to issue commands that involve several parts of the hand simultaneously. However, the aim of controlling a robotic device in grasping tasks allows a simplification in the description of finger kinematics by selecting a reduced set of parameters that could describe individual finger motion. One solution is to use a small number of synergies [11], [12]. Synergistic hand shaping would involve the movement of the digits in a highly coordinated, dependent pattern. Another low-dimension representation was reported in [13], where six vectors were used to represent finger motion during grasping of a variety of objects, for classification purposes.

In this paper, focusing on monkey grasping movements, we use the grasp *aperture* of a finger to approximately describe that finger's pose during the grasping task. Here we define grasp aperture of a given finger as the distance between the fingertip and the tip of the thumb. Using a

P. K. Artemiadis is with the Control Systems Lab, School of Mechanical Eng., National Technical University of Athens, 9 Heroon Polytechniou Str, Athens, 157 80, Greece partem@mail.ntua.gr

G. Shakhnarovich, and M. J. Black are with the Department of Computer Science, Brown University, Providence, RI 02912, USA {gregory, black}@cs.brown.edu

C. Vargas-Irwin, and J. P. Donoghue are with the Department of Neuroscience, Brown University, Providence, RI, 02912, USA, carlos_vargas_irwin@brown.edu, john_donoghue@brown.edu

state space model with hidden states to estimate the aperture from a small population of cells, we were able to accurately reconstruct the continuous variation of grasp apertures in a number of object grasping tasks.

II. Materials and methods

Our analysis used data previously collected during experiments where a monkey grasped moving objects. Below we briefly review the experimental setup, and proceed to describe the mathematical model used for decoding.

A. Experimental setup

One macaque monkey was implanted with one 96-electrode Bionic array (Cyberkinetics, Inc.) in the hand and arm area of primary motor cortex (rostral to the central sulcus at the level of the genu of the arcuate sulcus). Details of the array implantation and recording protocols are described elsewhere [14]. All procedures were in accordance with protocols approved by Brown University Institutional Animal Care and Use Committee. Neural signals were recorded using a Cerebus multichannel recording system (Cyberkinetics Inc.). Spike waveform recording was triggered using an amplitude threshold set to 4.5 RMS of the voltage values recorded for each channel. Single units were isolated offline using custom software implementing a template matching algorithm. Action potentials were then counted within fixed length, sliding time windows (bins) and the firing rate (number of spikes per unit time) within each bin was computed for each neuron. All decoding analysis was performed using these discrete approximations to the firing rate.

An optical motion capture system (Vicon-Peak, Inc.) was used for tracking hand and finger postures. Using six high-speed cameras operating at 120 frames per second, we tracked the position of multiple reflective markers (hemispheroids, 4mm in diameter) with sub-millimeter accuracy. The positions of between 10 and 19 markers attached to the monkey's hand were successfully recorded. The reflective marker placement at the joints of the monkey's right hand during motion capture is shown in Fig.1a. Marker positions and cortical signals were recorded simultaneously, and temporally aligned off-line. Labeled markers at individual frames are used to generate the three-dimensional (3D) reconstructions. In this paradigm, the grasp aperture of a finger is approximated by the distance between the most distal marker on the finger and the most distal marker of the thumb. This concept is illustrated in Fig. 1b.

The behavioral paradigm of the task entails a monkey intercepting and holding an object that swings towards it at the end of a string. After holding the object for one second the monkey is rewarded with fruit juice. The object rotates freely, and swings through various positions of the workspace. This task was designed to elicit the widest possible range of grasping movements, covering as much of the multidimensional space of hand motion as possible. Three different objects were used: a small ball (20mm in diameter), a cube (35×35×36mm) and a pipe (27mm in diameter, 167mm in length).

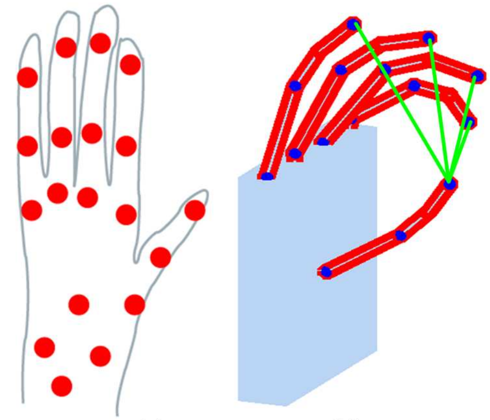

Fig. 1. a) Nineteen reflective markers are attached at the monkey's hand during motion capture. b) Aperture of each finger wrt the thumb.

B. State space model

Neural decoding of motor cortical activity involves estimating some movement-related variable from neural signals. Here our goal is to decode grasp aperture z_k for a particular finger at a particular time instance from the neural firing rates $\mathbf{u}_k$ of a population of cells. We take $\mathbf{u}_k = [u_{k,1}...u_{k,n}]^T$ to be a vector of firing rates for n cells at time instant k. A large number of algorithms have been previously used for decoding two- or three-dimensional hand kinematics [3], [7], [8], [9]. These previous methods typically embody the assumption that the firing rates are linearly related to the kinematic variables, either directly in the case of linear filters [8] or through a generative model in the case of Bayesian methods [9].

Unlike hand movements which have been well studied, the neural coding of finger aperture is not well understood. In particular, we do not know how aperture and firing rates are related (or even whether aperture is the relevant behavioral variable). Rather than assume a simple, direct, linear relationship, we adopt a more flexible decoding model in which we introduce "hidden", or "latent" variables we call $\mathbf{x}$. These hidden variables can, for example, represent the relationship between $\mathbf{u}$ and z in a higher dimensional space. The hope is that these hidden states model the unknown, intrinsic, movement parameters and relate these to the observed firing and finger aperture.

Specifically, we formulate the finger aperture decoding problem using the following state space model

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{L}\mathbf{u}_k + \mathbf{w}_k \quad (1)$$

$$z_k = \mathbf{C}\mathbf{x}_k + v_k. \quad (2)$$

Here., $\mathbf{x}_k \in \mathbb{R}^d$ is the hidden state vector at time instance kT, $k = 1, 2, \ldots$, (T being the sampling period), $\mathbf{u}_k \in \mathbb{R}^n$ is the vector of firing rates, and $z_k \in \mathbb{R}$ the finger aperture. The matrix $\mathbf{F}$ determines the dynamic behavior of the hidden state vector $\mathbf{x}$, $\mathbf{L}$ is a matrix relating firing rates $\mathbf{u}$ to the state vector $\mathbf{x}$, while $\mathbf{C}$ is a matrix that represents the relationship between the aperture z and the hidden states $\mathbf{x}$. $\mathbf{w}_k$ and v_k represent zero-mean Gaussian noise in the process and observation equations, respectively. The covariance matrix

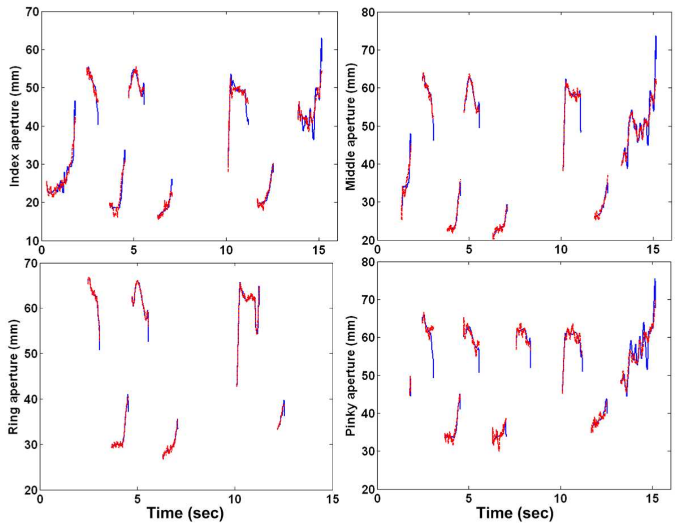

Fig. 2. Estimated (dashed red) and real (solid blue) aperture values for each finger. Time instances where markers are missing are not plotted. Note: nearly overlapping solid and dashed lines indicate high estimation accuracy.

of $\mathbf{w}_k$ is denoted by $\mathbf{Q} \in \mathbb{R}^{d \times d}$, and the variance of v_k is denoted by σ^2. I.e., $\mathbf{w}_k \sim N(\mathbf{0}, \mathbf{Q})$, $v_k \sim N(0, \sigma^2)$.

We note that this approach is more powerful than the previous Kalman filter methods for decoding hand trajectories from firing rates. The hidden state can represent internal (and hence unobserved) processes in the neural system. In future work we will explore whether the estimated hidden states can provide insight into the neural control of finger movement.

C. Model building and neural decoding

Fitting of the model requires estimation of the following parameters: the system matrices $\mathbf{F}, \mathbf{L}$ and $\mathbf{C}$, the state noise covariance matrix $\mathbf{Q}$, and the aperture (output) variance σ^2. Given a training set of length M, in which we observe both the system input $\mathbf{U}_M = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M]^T$ and system output $\mathbf{Z}_M = [z_1, z_2, \ldots, z_M]^T$, the parameters can be fit using an optimization procedure. Specifically, the objective is to minimize a quadratic prediction error on the training set. A standard approach is to use an iterative search algorithm [16]. A stability test of the predictor is performed to ensure that only models corresponding to stable predictors are tested. In the experiments reported below we used a standard implementation of this procedure in Matlab™ System Identification Toolbox.

Finally, the dimensionality d of the hidden state vector, called the model *order*, is a design parameter. It is selected by iterating through a range of values, so as to maximize the accuracy in prediction on the training set. In most of the cases tested, this procedure led us to select a 9^{th} order model.

Once the model has been trained, decoding aperture from neural signal is straightforward. Given the recorded $\mathbf{u}$, we estimate the hidden states $\mathbf{x}$ according to (1), and then recover the aperture z according to (2).

III. Results

Neural firing rates were computed from the recorded action potentials using a window of 400 ms, with sampling period T=1ms. Before building the decoding model, we first identified those neurons where the firing rate was most correlated with the aperture of a given finger. This led to the selection of $n = 8$ units. Thus, the input $\mathbf{u}$ to the model consisted of a vector of size eight, comprised of the firing rates of the selected units.

The aperture measurements of each finger were resampled offline at the frequency of the firing rates, using an anti-aliasing (lowpass) finite impulse response (FIR) filter. It must be noted that during motion of the hand, there were cases where at least one of the distal marker positions could not be computed, due to occlusion (and the subsequent failure of the optical system to track the marker). In these cases, the corresponding values for firing rates were ignored and not used during the model parameter estimation.

The model was fit using the iterative optimization algorithm as described in the previous section. The iterations were terminated when the number of iterations reached 200, or the expected improvement was less than 1%, or when a lower value of the criterion could not be found.

The method was initially tested in the ball catching scenario. Experimental measurements for a contiguous time period of 15 sec were acquired. During this scenario, the distal marker position of each finger was computed, and the four finger apertures were calculated. A separate model was trained for each finger's aperture. We divided the data into four segments, and performed four-fold cross validation, whereby three segments were used for training and the remaining segment for testing. We assess the accuracy of the decoding algorithm using correlation coefficient between the estimated (decoded) and the true finger aperture. In this experiment, the correlation coefficients for all four fingers were found equal to 0.99. Fig. 2 shows the estimated aperture values for each finger along with the ground truth (we show, for each segment, the prediction obtained in the fold where that segment was used for testing). Aperture values at time instances where markers were missing are not plotted.

A second set of experiments was conducted in order to test the decoding method in grasping two different objects (i.e., a cube and a pipe). We fit a single model to a training set constructed by combining 30 sec of data for each of the two objects. We then tested it on additional 5 sec segment for each object (not used in training). Only the aperture of the index finger was estimated in this experiment. Fig.3, 4 show the estimated aperture values for the index finger along with the ground truth. The correlation coefficient was 0.94 and 0.9 for the cube and pipe experiment respectively. The MSE of the estimates were 2.94mm and 3.44mm for the cube and pipe experiment respectively. The proposed model can decode neural activity of a population of neurons to derive continuous grasp aperture in cases where different objects are grasped.

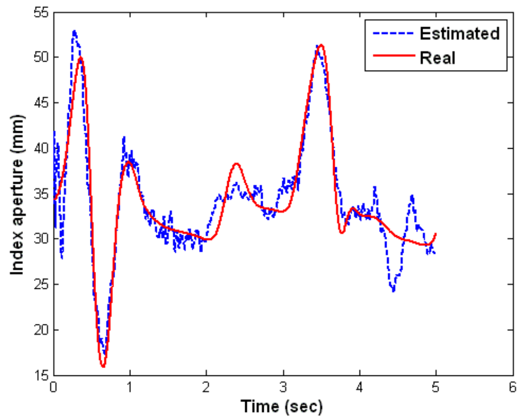

Fig. 3. Estimated and real index aperture values for cube grasp (test data).

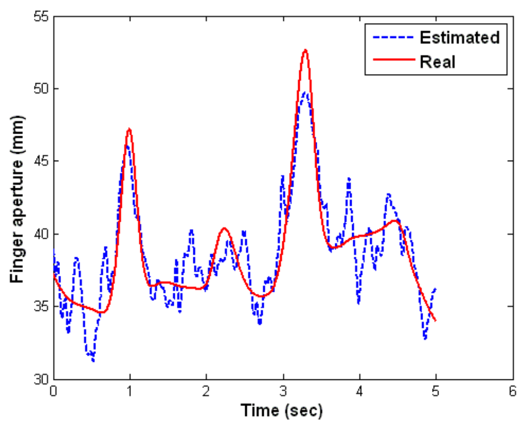

Fig. 4. Estimated and real index aperture values for pipe grasp (test data).

IV. CONCLUSIONS AND DISCUSSION

In this paper we have demonstrated decoding of grasp aperture from the activity of a population of neurons in motor cortex. To our knowledge, this is the first work in which this task is addressed. We see this as an important proof of concept, and potentially a step towards the goal of direct brain control of dexterous hand movements. Our results on reconstructing aperture values in an off-line experiment, using simultaneously recorded neural activity and 3D finger movement in three object grasping tasks, suggest that the task is feasible.

In the context of motor prosthetic applications, we see three important directions for further research, all of which we would like to pursue. One is the translation of hand control parameters (of which grasp aperture is one) to control signals for a specific robotic actuator. The second direction involves transferring the decoding framework from non-human primates to human clinical trials. Third, considering that natural grasping tasks involve arm reaching and wrist orientation in addition to grasp aperture, we will explore the simultaneous decoding of arm and wrist kinematics with grasping.

In this paper, we have successfully applied a state space model with latent variables; however, other methods developed recently for neural decoding have to be considered. These include linear filter models and Bayesian algorithms such as Kalman filters. In addition, we are investigating possible physiological semantics for the hidden states. Eventually our goal is the direct brain control of a high dimensional robotic hand prosthesis by paralyzed humans.

V. ACKNOWLEGMENTS

This work is partially supported by the European Neurobotics Program FP6-IST-001917, NIH-NINDS R01 NS 50867-01 as part of the NSF/NIH Collaborative Research in Computational Neuroscience Program, NINDS Javits Award NS25074, Center for Restorative and Regenerative Medicine, Rehabilitation R&D Service, Department of Veterans Affairs Medical Center, Providence, RI, and by the Office of Naval Research (award N0014-06-0185), . We also thank Kostas Kyriakopoulos and Patrick van der Smagt for their assistance.

REFERENCES

[1] Taylor, D., Helms Tillery, S., and Schwartz, A., ”Direct cortical control of 3D neuroprosthetic devices,” Science, 296, 1829-1832, 2002.

[2] Serruya, M. D., Hatsopoulos, N. G., Paninski, L., Fellows, M. R., and Donoghue, J. P. ”Brain-machine interface: Instant neural control of a movement signal,” Nature, (416):141-142, 2002.

[3] Wessberg, J., Stambaugh, C., Kralik, J., Beck, P., Laubach, M., Chapin, J., Kim, J., Biggs, S., Srinivasan, M., and Nicolelis, M., ”Real-time prediction of hand trajectory by ensembles of cortical neurons in primates,” Nature, 408:361-365, 2000.

[4] Carmena J.M., Lebedev M.A., Crist R.E., O’Doherty J.E., Santucci D.M., Dimitrov D.F., Patil P.G., Henriquez C.S., Nicolelis M.A., Learning to control a brain-machine interface for reaching and grasping by primates, Public Library of Science, vol. 1, pp. 1-16, 2003.

[5] Hochberg, L.R., Serruya, M.D, Friehs, G.M, Mukand, J.A., Saleh, M, Caplan, A.H., Branner, A., Chen, D., Penn, R.D., and Donoghue, J.P. ”Neuronal ensemble control of prosthetic devices by a human with tetraplegia,” Nature, 442, 164-171, 2006.

[6] Tao Kang, Jiping He and S. Helms Tillery. Determining natural arm configurations along a reaching trajectory, Experimental Brain Research, 167: 352-361, 2005

[7] Georgopoulos, A., Kalaska, J., Caminiti, R., and Massey, J., On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex, J. Neurosci., vol. 8, no. 11, pp. 1527-1537, 1982.

[8] Paninski, L., Fellows, M., Hatsopoulos, N., and Donoghue, J.P., Spatiotemporal tuning of motor cortical neurons for hand position and velocity, J. Neurophysiol., vol. 91, pp. 515-532, 2004.

[9] Wu, W., Black, M. J., Mumford, D., Gao, Y., Bienenstock, E., Donoghue, J. P., ”Modeling and decoding motor cortical activity using a switching Kalman filter,” IEEE Trans. Biomedical Engineering, 51(6):933-942, June 2004.

[10] Poliakov A. V., and Schieber M. H., ”Limited functional grouping of neurons in the motor cortex hand area during individuated finger movements: a cluster analysis,” J Neurophysiol 82: 3488-3505, 1999.

[11] d’Avella A., Saltiel P., and Bizzi E, ”Combinations of muscle synergies in the construction of a natural motor behavior,” Nat Neurosci 6:300-308, 2003.

[12] Mason C.R., Theverapperuma L.S., Hendrix C.M., and Ebner T.J., ”Monkey hand postural synergies during reach-to-grasp in the absence of vision of the hand and object,” J Neurophysiol 91:2826-2837, 2004.

[13] Theverapperuma L.S., Hendrix C.M., Mason C.R., and Ebner T.J., ”Finger movements during reach-to-grasp in the monkey: amplitude scaling of a temporal synergy,” Exp Brain Res. Mar;169(4):433-448. 2006

[14] Suner S., Fellows M.R., Vargas-Irwin C., Nakata K., and Donoghue J.P., ”Reliability of Signals from Chronically Implanted, Silicon-based Electrode Array in Non-human Primate Primary Motor Cortex,” IEEE Trans Neural Syst Rehabil Eng. 2005 Dec;13(4):524-41.

[15] Wu, W., Gao, Y., Bienenstock, E., Donoghue, J. P., Black, M. J., ”Bayesian population coding of motor cortical activity using a Kalman filter,” Neural Computation, 18:80-118, 2005.

[16] L. Ljung, ”System identification: Theory for the user,” Upper Saddle River, NJ: Prentice Hall, 1999.