Differences in Motor Cortical Representations of Kinematic Variables between Action Observation and Action Execution and Implications for Brain-Machine Interfaces

Francis R. Willett, Aaron J. Suminski, *Member, IEEE*, Andrew H. Fagg, *Member*, *IEEE*, and Nicholas G. Hatsopoulos

***Abstract*— Observing an action being performed and executing the same action cause similar patterns of neural activity to emerge in the primary motor cortex (MI). Previous work has shown that the neural activity evoked during action observation (AO) is informative as to both the kinematics and muscle activation patterns of the action being performed, although the neural activity recorded during action observation contains less information than the activity recorded during action execution (AE). In this study, we extend these results by comparing the representation of different kinematic variables in MI single /multi unit activity between AO and AE conditions in three rhesus macaques. We show that the representation of acceleration decreases more significantly than that of position and velocity in AO (population decoding performance for acceleration decreases more steeply, and fewer neurons in AO encode acceleration significantly as compared to AE). We discuss the relevance of these results to brain-machine interfaces that make use of neural activity during AO to initialize a mapping function between neural activity and motor commands.**

I. INTRODUCTION

The patterns of neural activity evoked in the primary motor cortex during action observation (watching another actor perform an action, AO) are similar to the patterns evoked during action execution (performing the action yourself, AE) and contain detailed information about the kinematics of observed movements and the muscle activation patterns that would cause them [3-6]. This fact has been exploited to initialize the mapping function between neural activity and motor commands in brain-machine interfaces for people who are paralyzed [1, 2]. While a paralyzed person can no longer perform motor acts, he or she can still observe, imagine, or attempt to make movements.

Previous results examining the relationship between single / multi units and observed actions reported that neural units contain less information about the motor task during AO and that this information is delayed relative to AE [3,4,6]. However, no systematic investigation has yet been made concerning more subtle differences in the *kinds* of motor variables represented in the primary motor cortex between AO and AE. Because there is no somatosensory feedback available during AO, we expect that the cortical representation of certain movement variables may decline more severely in AO relative to others. In this study, we show that the representation of hand acceleration in motor cortical activity decreases more steeply than that of hand position and hand velocity when rhesus macaques observe previously executed movements played back to them on a screen. The results, replicated in three monkeys and ten datasets, suggest that the representation of acceleration and force-related variables (hand acceleration, joint acceleration, joint torque) during AO is impoverished, and that therefore brain-machine interfaces decoding acceleration or force-related signals that expect to initialize a decoding function based on AO activity may perform poorly.

II. METHODS

A. Behavioral Task

Three adult rhesus macaques (*Macaca mulatta*) were trained to control a cursor in a two–dimensional workspace using a two-link robotic exoskeleton (BKIN Technologies, Kingston, ON). The animal sat in a primate chair with its arm abducted 90 degrees and supported by the robot such that all movements were made within the horizontal plane. Direct vision of the arm was precluded by a horizontal projection screen. Visual feedback was available via a cursor projected onto the screen. The position of the cursor was controlled by either the monkey's hand position or an automatic replay of the monkey's movements that were made earlier in the same experimental session (action observation).

The random target pursuit (RTP) task required the monkey to move a cursor (6 mm diameter) through a sequence of square targets (2.25cm^2), each of which was presented one at a time. Each time the monkey hit a target, a new target appeared immediately in a random location. In order to complete a successful trial the monkey was required to sequentially acquire seven targets (action execution) or watch seven targets being acquired (action observation). A trial was aborted if any movement between targets took

---

Manuscript received April 7, 2014. This work was supported by R01 NS048845 awarded from the NINDS and the NSF GRFP DGE-0951783.

F. R. Willett is with the Department of Biomedical Engineering at the Case Western Reserve University, the Dept. of Neurosciences at the Cleveland Clinic Lerner Research Institute, and the Louis Stokes Cleveland Veteran's Affairs Medical Center at Cleveland, OH, 44106 USA (frw9@case.edu).

A. J. Suminski is with the Department of Organismal Biology and Anatomy at the University of Chicago, Chicago, IL 60637 USA (asuminski@uchicago.edu).

A. H. Fagg is with the School of Computer Science at the University of Oklahoma, Norman, OK 73019 USA (fagg@cs.ou.edu).

N. G. Hatsopoulos is with the Department of Organismal Biology and Anatomy and the Committee on Computational Neuroscience at the University of Chicago, Chicago, IL 60637 USA (nicho@uchicago.edu).

longer than 2.5s or 5s (depending on the dataset), or if the monkey moved its arm outside of a 1 cm radius circle during action observation. Only successfully completed trials were included in the data analysis. An average of 347 seconds of action observation and 394 seconds of action execution were analyzed for each dataset. More details about these datasets can be found in previous reports [3, 6-8].

B. *Electrophysiology*

The monkeys were chronically implanted with a 100-electrode microelectrode array (Blackrock Microsystems, Inc., Salt Lake City, UT) in MI contralateral to the arm used for the task. During each recording session, signals from up to 96 electrodes were amplified (gain of 5000), band-pass filtered between 0.3 Hz and 7.5 kHz, and recorded digitally (14-bit) at 30 kHz per channel using a Cerebus acquisition system (Blackrock Microsystems, Inc., Salt Lake City, UT). The neural data used in our analyses were comprised of single and multiunit spiking events that were sorted offline for eight of the datasets and online using time-amplitude sorting windows for two. On average, 69.2, 71.6, and 58 single / multi units were sampled for monkey Mk, monkey B, and monkey Rs per session, respectively. All of the surgical and behavioral procedures were approved by the University of Chicago Institutional Animal Care and Use Committee and conform to the principles outlined in the *Guide for the Care and Use of Laboratory Animals*.

C. *Population Decoding*

All population decoding results used a linear model to predict kinematic variables at discrete, 50 ms time steps from a 600 ms history of binned firing rates (50 ms bin width). Specifically, signal $S_k(t)$ at discrete time bin t (where k is an index representing a kinematic variable) is reconstructed by the decoder as follows:

$$S_k(t) = \sum_{j=0}^{B-1} \sum_{i=0}^{C-1} A(k,j,i) * N(i,t-j) + c_k \ ,$$

where i indexes over the C neurons, j indexes over time bins, $N(i, t\text{-}j)$ is the spike count of neuron i at time bin $t\text{-}j$, A are the coefficients of the filter, and c_k is an offset parameter for signal S_k. Here, B = 12. The coefficients in A were fit using ridge regression as in [9]. The decoders were cross-validated using ten folds (i.e., decoder coefficients were trained on 90% of the data and tested on the remaining 10%; this process was repeated 10 times to get an unbiased estimate of the decoder's performance).

D. *Single / Multi Unit Encoding Model*

The following linear encoding model was used to determine which kinematic variables the neural units were significantly encoding:

$$f_n(t) = c_0 + c_{pos}*p(t+\tau_{pos}) + c_{vel}*v(t+\tau_{vel}) + c_{acc}*a(t+\tau_{acc}),$$

where $f_n(t)$ is the binned (50 ms bin width) firing rate of neuron n at discrete time step t, $p(t+\tau_{pos})$ is a column vector of X and Y hand position, $v(t+\tau_{vel})$ is a column vector of X and Y hand velocity, $a(t+\tau_{acc})$ is a column vector of X and Y hand acceleration, τ_{pos}, τ_{vel}, and τ_{acc} are time leads or lags representing a temporal offset of between -250 and 250 ms, c_{pos}, c_{vel}, and c_{acc} are row vectors of model coefficients and c_0 is a scalar offset term.

III. RESULTS

In this study, three rhesus macaques made 2D, planar reaching movements (action execution) and then watched these same movements played back to them as they held their arm stationary (action observation). The movements were completed one after the other in a continuous fashion with no defined breaks.

A. *Population Decoding*

To determine how well different kinematic variables were represented in MI during action observation (AO) vs. action execution (AE), we fit a linear decoding model to the data that predicted each kinematic variable as a function of the firing rates of neural units in the primary motor cortex. Our decoding model used binned firing rates (50 ms bin width) and 12 bins of firing rate history for each neuron (600 ms filter length) to predict kinematic variables binned at 50 ms time steps. We measured decoding performance by computing the 10-fold cross-validated fraction of variance accounted for (FVAF) by the decoder (similar to [9]), averaging over the different dimensions of the kinematic variables to compute a single FVAF score when appropriate (hand position, hand velocity, hand acceleration, and target position have both X and Y dimensions in our task). Decoders were trained and tested within the same condition (either AO or AE).

The overall decoding results are depicted in Figure 1. As expected, it is more difficult to decode most kinematic

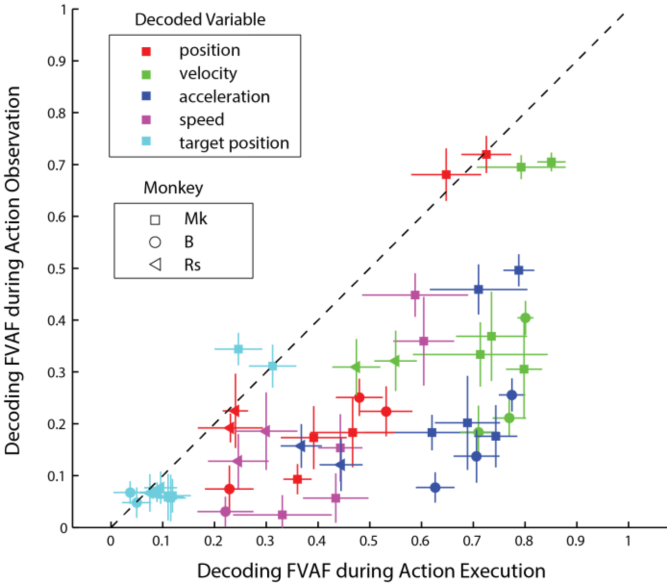

Figure 1. Average decoding performance (cross-validated, fraction of variance accounted for, FVAF) in action execution vs. action observation conditions depicted for five kinematic variables, ten datasets, and three monkeys. Lines protruding from the symbols represent 95% confidence intervals of the mean assuming that the data are normally distributed. For most datasets, decoding performance is superior during action execution, as expected.

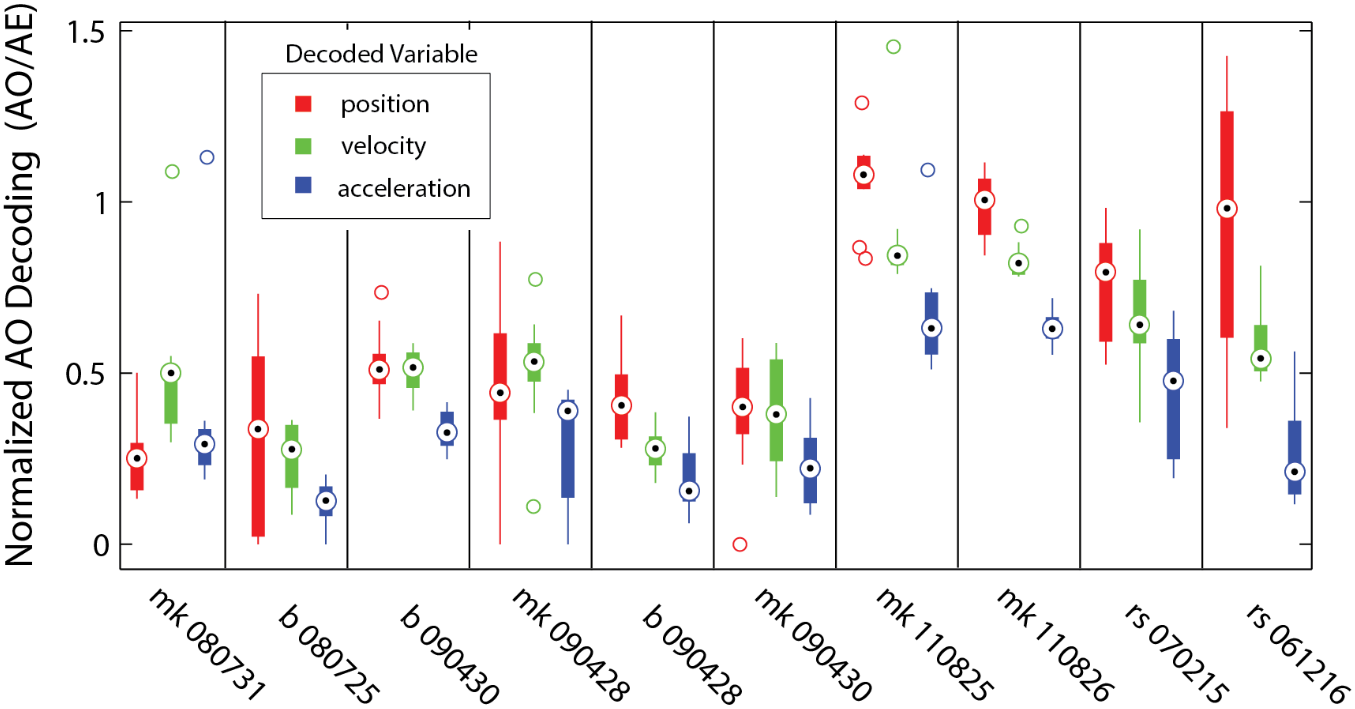

Figure 2. For each dataset, decoding performance for position, velocity, and acceleration in AO is compared to that in AE by dividing AO FVAF by AE FVAF to yield a normalized performance metric. The distributions of the 10-fold cross-validated performance metrics are drawn as boxplots (median values are plotted with black dots surrounded by circles; bottom and top of vertical bars represent first and third quartiles; outliers > 6 std from the mean are plotted as open circles). Most data points are below one because neural activity during AO is not as informative as neural activity during AE. For all ten datasets, acceleration decoding performance decreases more severely in AO as compared to velocity decoding. In all but one dataset (mk 080731), acceleration decoding performance also decreases more severely than position decoding. See Results for the statistics.

variables during AO as compared to AE. Velocity and acceleration have the strongest linear representation in the neural activity across both conditions (highest decoding FVAF), while position, target position, and speed are more poorly represented.

In Figure 2, we test whether or not the neural representation of certain variables decreases more severely than others in AO. Focusing on position, velocity, and acceleration, we divided AO decoding performance by AE decoding performance for each dataset and kinematic variable to reveal differences in AO decoding decline. Figure 2 shows that acceleration decoding performance decreases more steeply than that of velocity and position (statistically significant difference in 10 out of 10 datasets when compared to velocity, and in 9 out of 10 datasets when compared to position). Significance was determined by a t-test with $p<0.05$. To correct for multiple comparisons, we did a permutation test to compute the probability that, by chance, 19 or more t-tests would be significant for any one of the three variables when it was compared to the other two. We shuffled the normalized FVAF scores in Figure 2 within each dataset one million times and did the same t-tests. The largest number of significant t-tests was 9 (which occurred only twice), making the probability of observing 19 or more significant t-tests by chance at least less than 10^{-6}.

B. *Single Neural Unit Encoding*

In addition to population decoding, we also examined the encoding of kinematic variables at a single neuron level. We used methods similar to [10] to determine whether or not each neural unit was tuned to position, velocity, or acceleration in a way that could not be explained by correlations with other kinematic variables. For each neural unit, we fit and evaluated various linear encoding models that used one, 50 ms bin of X and Y hand position, X and Y hand velocity, and X and Y hand acceleration to predict the binned firing rate of the neural unit at each 50 ms time step (see Methods for the equation). We considered 11 different time leads and lags for each kinematic variable (evenly spaced between -250 and 250 ms), for a total of 11^3 possible "full" models that included all three kinematic variables.

For a neural unit to be significantly tuned to a kinematic variable v, the performance of the full encoding model containing v (measured by 10-fold cross-validated FVAF) had to be significantly greater (t-test, $p<0.05$) than that of an encoding model that did not include v but included the other two variables. This had to be true at *every* possible time lead and lag of the other two kinematic variables (11^2 t-tests must all be significant) but at *any* time lead or lag of v (11 possibilities to pass the set of 11^2 t-tests). A neural unit meeting this strict criterion could be said to encode the kinematic variable v in a way that could not be explained by encoding of a different kinematic variable that was correlated with v at a zero or non-zero temporal offset. Under this criterion, a neural unit could encode more than one variable. Our kinematic variable criterion is stricter than the typical requirement that a regression predicting firing rate from position, velocity, and acceleration be significant with $p < 0.05$ (436 of 677 units met the regression criterion but not the variable criterion, whereas 0 units met the variable criterion but not the regression criterion).

Figure 3 shows the results of this linear encoding analysis for each dataset, kinematic variable, and condition (AO vs. AE). The fraction of all tuned neurons (neurons that met the encoding criterion for *any* variable) that significantly encoded position, velocity, or acceleration during AE is plotted against the fraction observed during AO for each dataset and variable. For all ten datasets, the fraction of neural units encoding acceleration is greater in AE than AO, complementing the decoding results presented in section *A*. For 8 of the 10 datasets, there is also a greater fraction of velocity tuned units in AE relative to AO. The deficit in units encoding acceleration and velocity during AO is counterbalanced by a relative over-abundance of position-tuned neural units.

IV. DISCUSSION

In this study, we examined how the neural representation of kinematic variables in a 2D planar reaching task changes in MI during action observation (AO) relative to action execution (AE). While it is well known that movements are more weakly represented in AO relative to AE in general [3,4,6], it is unknown if the representation of certain movement variables decreases more severely than others during AO. In this study, we showed that not all variables decline equally; rather, the neural representation of hand acceleration declines more severely than that of hand position or hand velocity during AO. Both population decoding (Figure 1 and Figure 2) and neural unit encoding analyses (Figure 3) confirm this result in three monkeys and ten datasets. In most datasets (8 out of 10), the relative lack of acceleration-tuned neural units in AO was counterbalanced by an overabundance of position-tuned neural units (Figure 3).

Action observation has been proven to be a useful tool to initialize mapping functions between neurons and motor commands for brain-machine interfaces that are intended to be used with paralyzed people [1, 2]. An impoverished neural representation of acceleration could make it more difficult to implement a biomimetic brain-machine interface that aims to decode acceleration or force-like variables from primary motor cortex [11, 12]. In the ten datasets presented here, joint torque can be predicted from hand acceleration with a single bin linear model with very high accuracy (mean cross-validated FVAF of > 0.97 across all datasets), meaning that acceleration is highly correlated with joint torque and so the results likely apply to this variable as well.

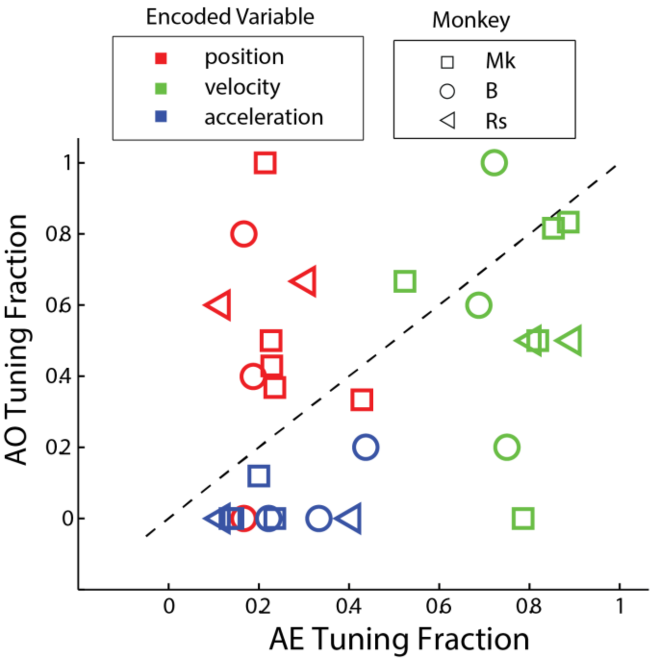

Figure 3. Fraction of all neural units tuned to at least one of the three variables (position, velocity, acceleration) that were tuned to position, velocity, or acceleration in a way that could not be explained by correlations with other kinematic variables. For all 10 datasets, more neural units were significantly tuned to acceleration during AE as compared to AO.

REFERENCES

[1] J. L. Collinger, B. Wodlinger, J. E. Downey, W. Wang, E. C. Tyler-Kabara, D. J. Weber, A. J. McMorland, M. Velliste, M. L. Boninger, and A. B. Schwartz, "High-performance neuroprosthetic control by an individual with tetraplegia," *The Lancet*, vol. 381, no. 9866, pp. 557–564, 16.

[2] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. van der Smagt, and J. P. Donoghue, "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, May 2012.

[3] D. Tkach, J. Reimer, and N. G. Hatsopoulos, "Congruent activity during action and action observation in motor cortex," *J. Neurosci.*, vol. 27, no. 48, pp. 13241–13250, Nov. 2007.

[4] J. Dushanova and J. Donoghue, "Neurons in primary motor cortex engaged during action observation," *Eur. J. Neurosci.*, vol. 31, no. 2, pp. 386–398, Jan. 2010.

[5] K. Alaerts, S. P. Swinnen, and N. Wenderoth, "Is the human primary motor cortex activated by muscular or direction-dependent features of observed movements?," *Cortex*, vol. 45, no. 10, pp. 1148–1155, Dec. 2009.

[6] A. J. Suminski, D. C. Tkach, and N. G. Hatsopoulos, "Exploiting multiple sensory modalities in brain-machine interfaces," *Neural Netw*, vol. 22, no. 9, pp. 1224–1234, Nov. 2009.

[7] A. J. Suminski, D. C. Tkach, A. H. Fagg, and N. G. Hatsopoulos, "Incorporating feedback from multiple sensory modalities enhances brain-machine interface control," *J. Neurosci.*, vol. 30, no. 50, pp. 16777–16787, Dec. 2010.

[8] F. R. Willett, A. J. Suminski, A. H. Fagg, and N. G. Hatsopoulos, "Improving brain–machine interface performance by decoding intended future movements," *J. Neural Eng.*, vol. 10, no. 2, p. 026011, Apr. 2013.

[9] A. H. Fagg, G. W. Ojakangas, L. E. Miller, and N. G. Hatsopoulos, "Kinetic Trajectory Decoding Using Motor Cortical Ensembles," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 17, no. 5, pp. 487–496, Oct. 2009.

[10] E. Stark, R. Drori, I. Asher, Y. Ben-Shaul, and M. Abeles, "Distinct movement parameters are represented by different neurons in the motor cortex," *European Journal of Neuroscience*, vol. 26, no. 4, pp. 1055–1066, 2007.

[11] A. J. Suminski, F. R. Willett, A. H. Fagg, M. Bodenhamer, and N. G. Hatsopoulos, "Continuous decoding of intended movements with a hybrid kinetic and kinematic brain machine interface," *Conf Proc IEEE Eng Med Biol Soc*, vol. 2011, pp. 5802–5806, 2011.

[12] P. Y. Chhatbar and J. T. Francis, "Towards a naturalistic brain-machine interface: hybrid torque and position control allows generalization to novel dynamics," *PLoS ONE*, vol. 8, no. 1, p. e52286, 2013.