

## Real-time Fusion of Gaze and EMG for a Reaching Neuroprosthesis

Elaine A. Corbett, *Student Member, IEEE*, Konrad P. Kording, and Eric J. Perreault, *Member, IEEE*

**Abstract**—For rehabilitative devices to restore functional movement to paralyzed individuals, user intent must be determined from signals that remain under voluntary control. Tracking eye movements is a natural way to learn about an intended reach target and, when combined with just a small set of electromyograms (EMGs) in a probabilistic mixture model, can reliably generate accurate trajectories even when the target information is uncertain. To experimentally assess the effectiveness of our algorithm in closed-loop control, we developed a robotic system to simulate a reaching neuroprosthetic. Incorporating target information by tracking subjects' gaze greatly improved performance when the set of EMGs was most limited. In addition we found that online performance was better than predicted by the offline accuracy of the training data. By enhancing the trajectory model with target information the decoder relied less on neural control signals, reducing the burden on the user.

### I. INTRODUCTION

FOR people whose arms have been paralyzed by spinal cord injury (SCI) movement can be restored using a neuroprosthetic known as Functional Electrical Stimulation (FES), where the spinal cord is effectively by-passed and the peripheral nerves are stimulated electrically [1], or other robotic assistive devices [2]. To date, most implementations of full arm reaching have involved pre-programmed patterns of activation, controlled by switching mechanisms through respiration [3] or contra-lateral shoulder movement [4], resulting in unnatural control of limited movement patterns. To allow a more flexible interface with multiple degrees of freedom (DOFs), one of the most challenging problems is the determination of user intent from the physiological signals that remain under voluntary control. At high levels of cervical SCI, individuals have no control over a majority of the arm muscles, thus inferring free reaches through residual movement or electromyograms (EMG) alone is not feasible.

A number of groups in the brain-machine-interface (BMI) field have shown that reconstruction of reach trajectories from neural signals can be greatly improved when infor-

mation about the reach target is available [5–8]. We have previously shown that reaching movements can be effectively reconstructed when a single shoulder EMG is combined with target information obtained by tracking eye movements [9], [10]. Eye movements are a natural, unobtrusive way to learn about a subject's intentions. However, the target information obtained will be uncertain; though people almost always look at a target before reaching to it, they may also gaze at other locations. We have shown that probabilistic mixture models can account for this and accurately reconstruct reaching movements even when there is high uncertainty about the target [11].

In many cases decoding algorithms have been tested by recording the neural signals as natural arm movements are made, and then performing offline evaluations of how well those movements can be reconstructed. However, researchers have recently shown that offline accuracy does not necessarily predict online performance [12], [13]. This is problematic as many proposed decoding algorithms have not been tested in closed-loop. Additionally, in a real neuroprosthetic system there will be no natural reaches available to train the models.

We have developed a robotic system to simulate a neuroprosthesis, thus providing a means to evaluate our approach to combining EMG and gaze information in closed-loop control. This simulation was carried out by having a robot move a subject's arm throughout a reaching workspace based on the decoded velocity and position. We found that incorporating the gaze data produced dramatic improvements in control with a single shoulder EMG channel.

### II. METHODS

#### A. Decoding Algorithms

We compared two decoding approaches: a simple decoder that used only the EMGs as inputs and another that incorporated target information. We tested the latter both with perfect target information and also using information found from tracking subjects' eye movements. For both algorithms we employed the Kalman filter (KF) framework for decoding, assuming linear dynamics and Gaussian noise:

$$\mathbf{x}_t = [\mathbf{z}_t \dot{\mathbf{z}}_t \ddot{\mathbf{z}}_t]^T = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{w}_t, \quad (1)$$

where  $\mathbf{x}_t$  is the state vector at time  $t$ ,  $\mathbf{z}_t \in \mathbb{R}^p$  represents the hand position,  $\mathbf{w}_t$  is the process noise with  $p(\mathbf{w}) \sim N(0, \mathbf{Q})$ , and  $\mathbf{Q}$  is the state covariance matrix.

To create a directional trajectory model, we added the target position to the state space (KFT), thereby linearly incorporating it into the trajectory model [7], [8]:

$$\mathbf{x}_t = [\mathbf{z}_t \dot{\mathbf{z}}_t \ddot{\mathbf{z}}_t \mathbf{z}T_t]^T = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{w}_t, \quad (2)$$

where  $\mathbf{z}T_t \in \mathbb{R}^p$  is the vector of target positions. In all cases the observation model was considered to be linear, with

Manuscript received March 29, 2012. This work was supported by the NSF program in Cyber-Physical Systems (NSF/CNS 0939963).

E. A. Corbett is with the Department of Biomedical Engineering, Northwestern University, Evanston, IL 60208 USA (phone: 312-238-5828; fax: 312-238-2208; e-mail: [ecorbett@u.northwestern.edu](mailto:ecorbett@u.northwestern.edu)).

K. P. Kording is with the Department of Physiology and the Department of Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL 60611 USA (email: [kk@northwestern.edu](mailto:kk@northwestern.edu)).

E. J. Perreault is with the Department of Biomedical Engineering and the Department of Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL 60611 USA (email: [e-perreault@northwestern.edu](mailto:e-perreault@northwestern.edu)).

Gaussian noise. Observations were generated from the corresponding window of EMG by extracting two features from each channel – the RMS value and the number of zero crossings (above a threshold), a frequency-related feature. The square root transform of these features was taken to obtain more Gaussian-like distributions.

For the case when the target estimates were based on eye movements, multiple potential targets needed to be considered. To achieve this we used a probabilistic mixture model (mKFT) over each of the potential targets [5]. The KFT recursion was performed for each possible target,  $\mathbf{zT}$ , and a weighted sum of the outputs was taken. The weights were proportional to the prior for that target, and the likelihood of the model given that target. The weights were thus initialized based on the gaze data and they converged to the most likely trajectory as the neural information was integrated over the course of the reach. Further details of the algorithms can be found in [9–11].

### B. Experimental Setup and Protocols

Six able-bodied subjects (3 male, 3 female) participated in the experiments. Each subject provided informed written consent to a protocol approved by Northwestern University's Institutional Review Board. All subjects were right-arm dominant.

A retractable stylus was attached to the handle of a 3 DOF HapticMaster robot. Subjects were seated comfortably facing two touch-screen monitors that were at different distances to the subject in the Z direction (Fig. 1a). The subject held the robot handle which was positioned directly in front of them; the robot then moved their right arm forward through a reach so that the stylus would touch the monitors when the arm was extended. The HapticMaster velocity was controlled at 60Hz, with low-gain PID feedback on the position error to maintain positional fidelity.

The EMG signals were anti-alias and band-pass filtered and recorded at 2400Hz. The monitor, HapticMaster and head positions were recorded at 60Hz using an Optotrak motion analysis system so that gaze data and positions on the monitors could be transformed into the HapticMaster workspace. We recorded eye movements with an ASL EYETRAC-6 head-mounted eye tracker. All signals were recorded simultaneously and processed at 60Hz, thus the EMGs were divided into 16ms windows for feature extraction.

We tested all three algorithms (KF, KFT and mKFT) with two sets of EMGs, in an attempt to simulate the signals that would be available at different levels of SCI. To simulate an injury below the fourth cervical level (C4) we used just the

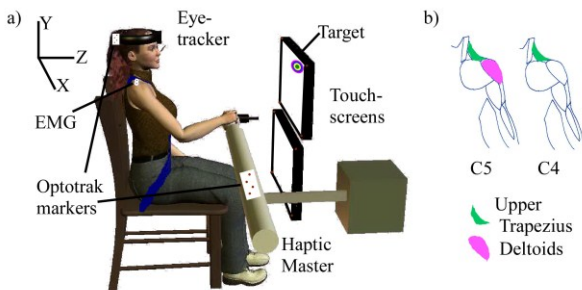


Fig. 1 a) Experimental setup; b) EMG simulation of SCI levels.

upper trapezius and for C5 we also included the anterior, middle and posterior deltoids (Fig. 1b). Each subject performed three experimental sessions: one without target information where the KF was tested at both simulated injury levels (in separate blocks); one where the KFT and mKFT (perfect target information and eye-tracking) were tested at C5; a third where the KFT and mKFT were tested at C4. The orders of the algorithms and simulated injury levels were randomized across subjects.

### C. Training the Decoders

Training data was required to estimate the decoder parameters. Because we wanted control to be intuitive, it was important that the EMGs controlling the decoder correspond as closely as possible to natural reaches. However, as the target population would be unable to generate unassisted reaches, it made sense to have the robot move along an "ideal" trajectory (linear in the kinematics and target) as the subject attempted to move along with the reach.

During training, 18 targets spanning the reachable area of the two monitors each appeared twice in random order. The reach began with the HapticMaster in the original starting position and, after an audible go cue, subjects performed the reach towards the target while EMGs were recorded (Fig. 2). During the reach the subject was instructed to hold on and gently assist the movement. The reach was ended when stylus tip reached the touch-screen, at which time the robot returned to the start position.

For both models the parameters  $A$ ,  $Q$  and the observation model and covariance matrices were estimated from training data using the maximum likelihood solution. In the case of the KFT, the final recorded position of the stylus was appended to the state vector for training, taking the place of the target estimate.

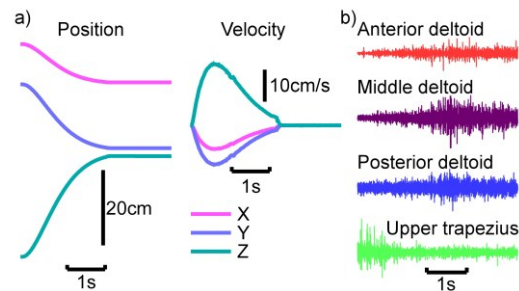


Fig. 2. Sample training reach a) kinematics and b) EMG

### D. Real-time Decoder Evaluation

After the models had been trained they were evaluated in a target acquisition task. For each trial a target randomly appeared, 1s before the go cue. The goal was to place the stylus in center of the target. Reaches were initiated when any EMG channel doubled relative to its level prior to the go cue. For the C4 level, the contralateral upper trapezius was also recorded to allow subjects to initiate reaches where they would not normally activate the ipsilateral muscle, by shrugging their left shoulder. However, it was not included as a part of the decoder (Fig. 3). After initiation, the decoded velocity and position were used to control the HapticMaster.

When testing the KF, only the EMGs were used as inputs to the decoder. In the case of the KFT, the state vector was initialized with the actual location of the target. For the

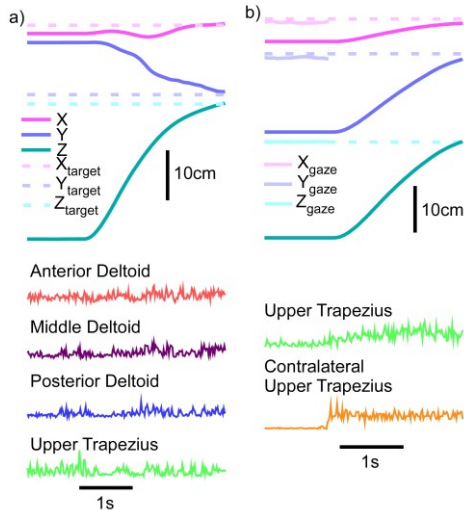


Fig. 3. Sample testing reaches, kinematics and square-root transformed RMS of EMG a) KF at C5 and b) mKFT at C4

mKFT, the gaze data from the one-second period prior to initiation was used to estimate three potential targets with which to initialize a corresponding mixture component (Fig. 3b). The three-dimensional location of the eye gaze was first calculated by projecting its direction onto the monitors. The first, middle and last samples were selected, and all other samples were assigned to a group according to which of the three was closest. The means of these three groups were used to initialize three KFTs in the mixture model and their priors were assigned proportional to the number of samples in them. If the subject looked at multiple positions prior to reaching, this method ensured with a high probability that the correct target was accounted for by one of the filters in the mixture.

After sufficient practice to produce a learning plateau, each subject performed forty test reaches for each decoding model. Control performance was evaluated based on the position of the stylus at the end of the reach. The target acquisition rate and final distance to target were calculated. In addition, the proportion of target variance accounted for (VAF) was calculated separately in the X and Y directions, by normalizing the final distance to the target by the total variance of the targets. As there was greater target variance in the Y direction, this allowed us to quantify the level of control in the two dimensions. As all of the targets were on the two monitors their positions were roughly similar in the Z dimension.

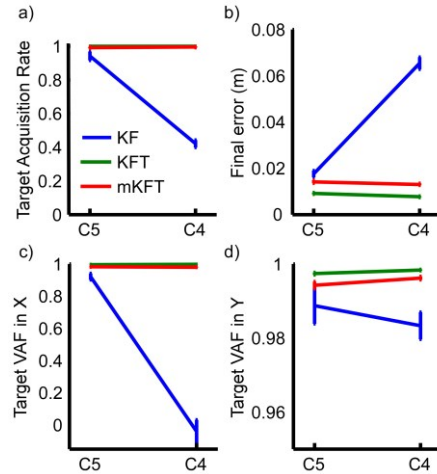
Finally the online  $R^2$  was evaluated between the performed reach and the "ideal" reach simulated between the start position and the target using the method for generating training reaches. This metric was used for comparison with the  $R^2$  of the training data, which was calculated using leave-one-out cross-validation to compare the EMG (and final target in the case of the KFT) based offline reconstructions to the training reaches.

### III. RESULTS

#### A. Control Performance

Unsurprisingly, when the decoder was given perfect target information (KFT) the reaches were very accurate, regardless of the quantity of EMG available (Fig. 4). Performance

with the eye-tracking (mKFT) was only slightly less accurate and also remained consistent across the two simulated injury levels. For the KF, with only EMG as an input, the simulated level of injury naturally had a large effect. The target acquisition rate at C5 with all four EMGs was close to 1, whereas at C4 it dropped to 40%. The average target error was roughly 5cm greater than that of the mKFT. However, this error was effectively determined by the distribution of targets in the workspace. As the upper trapezius was primarily activated during movements in the positive Y direction subjects were able to accurately control this dimension, but were unable to move in the X direction. This is illustrated in the target VAF, which is 0 in X and close to 1 in Y (Fig. 4c-d).



#### B. Comparison of Offline Accuracy and Performance

A central question in decoding research is how online and offline performance correspond. We found that the online  $R^2$  was consistently higher than the training  $R^2$  (Fig. 5). Even though subjects were in no way constrained to follow the "ideal" reach path, it was more closely replicated online than in reconstructing the training reaches where that path had been enforced. In particular, the KF at C4 produced a wide range of training accuracies, as the extent to which the upper trapezius was activated during training varied across subjects. However, all subjects were able to perform highly accurate control in the Y-direction, as reflected by the consistently high  $R^2$  values. As mentioned above, there was little control in the X-direction – the high  $R^2$  values are due to the

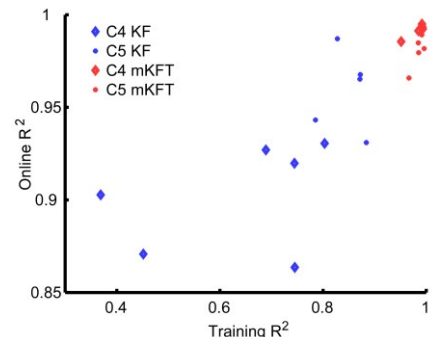


Fig 5. Online  $R^2$  plotted against  $R^2$  of training data for the different algorithm and simulated injury level combinations for each subject.

majority of the reach variance being in the Y and Z-directions.

#### IV. DISCUSSION

For people who have sustained a high-level SCI, FES control of reaching is a challenging task. The available proximal muscles are not sufficient to provide effective control; the incorporation of additional peripheral sensors, such as eye-trackers, is therefore an obvious solution. Adding target information into the trajectory model is an intuitive way to enhance control; the target informs us about the reach dynamics, resulting in a more well-defined trajectory. As our training reaches were generated from a stereotyped trajectory model as opposed to natural reaches, the uncertainty about the model was further minimized. This served to decrease the reliance on neural control signals, requiring less effort from the user.

To effectively evaluate signal sources and algorithm approaches it is essential that they be compared in closed-loop control. While there are many examples in the literature of closed-loop systems with non-human primates, usually controlling a virtual interface, few of them have compared algorithms explicitly. A small number of human experiments have compared algorithms [15], or signal sources [14] in closed-loop. Cunningham et al used a closed-loop BMI simulator to explicitly compare offline and online performance while changing the bin-width of a decoder, finding that online and offline evaluations produced strikingly different results [13]. While a larger bin-width provided smoothness in offline reconstructions, smaller bins are more useful online as they allow faster response times for subjects' error corrections. We did not vary any parameter to explicitly influence offline accuracy, however we did compare how well the "ideal" trajectory was replicated both on- and offline.

While in this study both offline and online accuracy with target information were very high, we found that KF performance without target information, when the subjects could interact with the decoder, was dramatically better than the offline accuracy predicted. This means that the improvements found by incorporating target information are less than indicated from previous offline studies [9], [10]. However, in the most severe cases, simulated here in the C4 level, the target information was necessary to produce functional movements in all directions. Furthermore, while accurate reaches were performed in the C5 case using only EMG, it often required more effort from the user. Qualitatively, subjects reported that the decoders incorporating target information were much easier to control than those without; the model with perfect target information often seemed effortless. This is somewhat illustrated by the fact that the online  $R^2$  is lower for the C5 KF versus mKFT (Fig. 5). While the KF reaches were almost as accurate at the target (Fig. 4), the trajectories produced were less consistent.

Gaze information is extremely useful when predicting desired trajectories, and we have shown here that it is effective and practical for use in a real-time system. Target information could equally be obtained by other means such as intracortical recordings or scanning the workspace. Even a small quantity of neural information can compensate for uncertainty in the target information; a single EMG combined with the gaze here enabled accurate control. Further-

more, even when there was sufficient neural data for effective control the target information helped generate more "ideal" trajectories, relying less on the users' neural control and thereby reducing their cognitive burden.

#### ACKNOWLEDGMENT

The authors thank Tim Haswell and Ben Walker for their work developing the data acquisition and robot control systems, and Dr Nicholas Sachs for experimental assistance.

#### REFERENCES

- [1] D. T. Yu, R. F. Kirsch, A. M. Bryden, W. D. Memberg, and A. M. Acosta, "A neuroprosthesis for high tetraplegia," *J Spinal Cord Med*, vol. 24, no. 2, pp. 109–113, 2001.
- [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] B. T. Smith, M. J. Mulcahey, and R. R. Betz, "Development of an upper extremity FES system for individuals with C4 tetraplegia," *IEEE Transactions on Rehabilitation Engineering*, vol. 4, no. 4, pp. 264–270, 1996.
- [4] N. Hoshimiya, A. Naito, M. Yajima, and Y. Handa, "A multichannel FES system for the restoration of motor functions in high spinal cord injury patients: a respiration-controlled system for multijoint upper extremity," *IEEE Transactions on Biomedical Engineering*, vol. 36, no. 7, pp. 754–760, 1989.
- [5] B. M. Yu, C. Kemere, G. Santhanam, A. Afshar, S. I. Ryu, T. H. Meng, M. Sahani, and K. V. Shenoy, "Mixture of trajectory models for neural decoding of goal-directed movements," *Journal of neurophysiology*, vol. 97, no. 5, p. 3763, 2007.
- [6] L. Srinivasan, U. T. Eden, A. S. Willsky, and E. N. Brown, "A state-space analysis for reconstruction of goal-directed movements using neural signals," *Neural computation*, vol. 18, no. 10, pp. 2465–2494, 2006.
- [7] C. Kemere and T. Meng, "Optimal estimation of feed-forward-controlled linear systems," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings (ICASSP'05)*, 2005, vol. 5.
- [8] G. H. Mulliken, S. Musallam, and R. A. Andersen, "Decoding trajectories from posterior parietal cortex ensembles," *Journal of Neuroscience*, vol. 28, no. 48, p. 12913, 2008.
- [9] E. A. Corbett, E. J. Perreault, and K. P. Kording, "Decoding with limited neural data: a mixture of time-warped trajectory models for directional reaches," *Journal of Neural Engineering*, forthcoming.
- [10] Corbett, E.A., Perreault, E.J., and Kording, K.P., "Mixture of time-warped trajectory models for movement decoding," in *Advances in Neural Information Processing Systems*, 2010, vol. 23.
- [11] E. A. Corbett, N. A. Sachs, K. Kording, and E. J. Perreault, "Dealing with noisy gaze information for a target-dependent neural decoder," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 5428–5431.
- [12] S. M. Chase, A. B. Schwartz, and R. E. Kass, "Bias, optimal linear estimation, and the differences between open-loop simulation and closed-loop performance of spiking-based brain-computer interface algorithms," *Neural Networks*, vol. 22, no. 9, pp. 1203–1213, 2009.
- [13] J. P. Cunningham, P. Nuyujukian, V. Gilja, C. A. Chestek, S. I. Ryu, and K. V. Shenoy, "A closed-loop human simulator for investigating the role of feedback control in brain-machine interfaces," *Journal of Neurophysiology*, vol. 105, no. 4, pp. 1932–1949, 2011.
- [14] E. A. Corbett, E. J. Perreault, and T. A. Kuiken, "Comparison of electromyography and force as interfaces for prosthetic control," *J Rehabil Res Dev*, vol. 48, no. 6, pp. 629–42, 2011.
- [15] S.-P. Kim, J. D. Simeral, L. R. Hochberg, J. P. Donoghue, and M. J. Black, "Neural control of computer cursor velocity by decoding motor cortical spiking activity in humans with tetraplegia," *J. Neural Eng.*, vol. 5, no. 4, pp. 455–476, Dec. 2008.