33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Dealing with Noisy Gaze Information for a Target-dependent Neural Decoder

Elaine A. Corbett, Nicholas A. Sachs, *Member, IEEE*, Konrad P. Körding, and Eric J. Perreault, *Member, IEEE*

***Abstract* — We tend to look at targets prior to moving our hand towards them. This means that our eye movements contain information about the movements we are planning to make. This information has been shown to be useful in the context of decoding of movement intent from neural signals. However, this is complicated by the fact that occasionally, subjects may want to move towards targets that have not been foveated, or may be distracted and temporarily look away from the intended target. We have previously accounted for this uncertainty using a probabilistic mixture over targets, where the gaze information is used to identify target candidates. Here we evaluate how the accuracy of prior target information influences decoding accuracy. We also consider a mixture model where we assume that the target may be foveated or, alternatively, that the target may not be foveated. We found that errors due to inaccurate target information were reduced by including a generic model representing movements to all targets into the mixture.**

I. Introduction

Eye movements can provide us with a huge amount of information about intended arm movements, since people usually look at a target before reaching to it [1]. This information could be useful for decoding intent, as needed for a range of rehabilitation applications including restoring movement control to individuals who have lost the use of a limb to paralysis. Eye tracking is often used to assist this population with computer interactions [2]. However, the success of these devices is limited by the fact that gaze alone is a problematic input signal. It can be challenging to determine which saccades are intended as control signals, and it is quite difficult to control eye movements precisely for an extended period of time.

When the goal is to restore arm motion through electrical stimulation or the assistance of a robotic device, it is even more critical that saccades do not generate unintentional movements. Furthermore, the patient's ability to look around their workspace should not be restricted. It is therefore likely that gaze information would be most useful when combined with physiological signals under the user's voluntary control [3].

A number of groups have shown that neural decoding can be greatly improved by taking advantage of the directional nature of reaching [4-7]. Following this logic, we have previously used target estimates from gaze information to enhance our model of the reach trajectory. Since a person could saccade to other locations in addition to the target during the period before a reach, we needed to account for uncertainty in the target estimates. We performed a mixture model, allowing for a probabilistic distribution of targets. This approach greatly improved our decoding of arm movements from electromyograms (EMG) [3].

Here we considered what would happen as the quality of our target estimates declined. While eye movements are a powerful source of information about a patient's intended reach, it is imperative that the system be able to cope in the face of erroneous target estimates. A decoder that relies too heavily on the gaze data could potentially produce unwanted movements and be unsafe. We examined the effectiveness of the mixture model and its dependence on the availability of neural data. We also developed a solution to mitigate the worst-case-scenario that occurs when the correct target is not foveated. We defined a generic model which was our best estimate of the movement when the target was unknown, and incorporated it into the mixture. When only inaccurate target estimates are available, our solution will converge to that of the generic model based on the evidence from the EMG.

II. Methods

A. Decoding Framework

We evaluated our approach at decoding human movement over a wide range of dynamics. We recorded arm kinematics, EMGs and eye movements of able-bodied subjects during unconstrained 3D reaches to targets over a large workspace. Four subjects performed reaches at varying speeds, as they would in everyday life. Subjects were seated as they reached towards 16 LEDs in blocks of 150s, which

Manuscript received April 15, 2011. This work was supported in part by the NSF program in Cyber-Physical Systems (NSF/CNS 0939963), and in part with federal funds from the National Center for Research Resources (NCRR), National Institutes of Health (NIH), through the Clinical and Translational Science Awards Program (CTSA), a trademark of DHHS, part of the Roadmap Initiative, "Re-Engineering the Clinical Research Enterprise," Northwestern University UL1RR0254741. Content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

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).

N. A. Sachs is with the Department of Biomedical Engineering, Northwestern University, Evanston, IL 60208 USA (e-mail: nsachs@northwestern.edu).

K. P. Körding is with Department of Physiology and the Department of Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL 60611 USA (email: 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).

were located on two planes positioned such that all targets were just reachable (Fig 1). The target LED was lit for one second prior to an auditory go cue, at which time the subject would reach to the target. An approximate total of 450 reaches were performed per subject. The subjects provided informed consent, and the protocol was approved by the Northwestern University Institutional Review Board. EMG signals were measured from the ipsilateral brachioradialis, biceps, two triceps, pectoralis major, and the three deltoid and upper trapezius muscles of the shoulder.

The EMG signals were band-pass filtered between 10 and 1,000 Hz, and subsequently anti-aliased filtered. Hand, wrist, shoulder and head positions were tracked at 60Hz using an Optotrak motion analysis system. We simultaneously recorded eye movements at 60Hz with an ASL EYETRAC-6 head mounted eye tracker.

100 reaches were assigned to the test set for each subject, and the rest were used for training. As the state we used hand positions and joint angles (3 shoulder, 2 elbow, position, velocity and acceleration, 24 dimensions). Joint angles were calculated from the shoulder and wrist marker data using digitized bony landmarks which defined a coordinate system for the upper limb as detailed by Wu et al. [8]. As the motion data were sampled at 60Hz, the square root of the mean absolute value of the EMG in the corresponding 16.7ms windows was used as an observation of the state at each time-step. We used the square root as it produced a more Gaussian-like distribution.

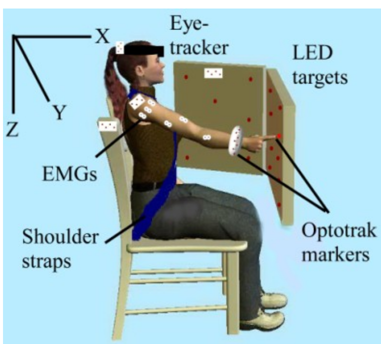

Fig. 1. Experimental setup.

B. *Mixture of Targets Model*

We employed the standard Kalman filter (KF) framework for decoding [9,10], assuming linear dynamics and Gaussian noise. We call this the generic KF trajectory model:

$$\mathbf{x_t} = [x_t\, \dot{x}_t\, \ddot{x}_t]^{\mathrm{T}} = A\mathbf{x_{t-1}} + \boldsymbol{w_t}, \tag{1}$$

where $\mathbf{x}$ is the state vector, $\boldsymbol{x_t} \in \mathbb{R}^{\boldsymbol{p}}$ represents the hand and joint angle positions, $\boldsymbol{w}$ is the process noise with $\mathrm{p}(\boldsymbol{w}) \sim N(0,Q)$, and Q is the state covariance matrix.

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

$$\mathbf{x_t} = [x_t\, \dot{x}_t\, \ddot{x}_t\;\, \boldsymbol{xT_t}]^{\mathrm{T}} = A\mathbf{x_{t-1}} + \boldsymbol{w_t}, \tag{2}$$

where $\boldsymbol{xT_t} \in \mathbb{R}^{\boldsymbol{g}}$ is the vector of target positions, with dimensionality less than or equal to that of $\boldsymbol{x_t}$. We also used a linear Gaussian observation model.

To account for the inevitable uncertainty in our eye-movement based target estimates, we drew on the recent literature [4,11] and used a probabilistic mixture model (mKFT) over each of the N potential targets T:

$$P(\mathbf{x_t}|\boldsymbol{y_{1\ldots t}}) = \sum_{T=\mathrm{T1}}^{\mathrm{TN}} P(\mathbf{x_t}|\boldsymbol{y_{1\ldots t}}, \boldsymbol{xT_t})P(\boldsymbol{xT_t}|\boldsymbol{y_{1\ldots t}}) \tag{3}$$

where $\boldsymbol{y_t}$ is the observation vector. We perform the KFT recursion for each possible target, xT, and our solution is a weighted sum of the outputs. The weights are proportional to the prior for that target (found from the gaze data), and the likelihood of the EMG given that target [3]. Thus, if the EMG data provides strong evidence in favor of one of the targets its weight will dominate and the prediction will quickly converge to that model.

As the targets were situated on two planes, the three-dimensional location of the eye gaze was found by projecting its direction onto those planes. The first, middle and last eye samples for the second preceding the reach were selected as potential targets, and all other samples were assigned to a group according to which of the three was closest. The mean and variance of these three groups were used to initialize three KFTs in the mixture model. The priors of the three groups were assigned proportional to the number of samples in them. This procedure assumed that the subject was looking close to the correct target during at least one of the three selected times prior to movement. If the subject looks at multiple positions prior to reaching, this method ensures with a high probability that the correct target was accounted for in one of the filters in the mixture.

We compared the generic KF (with no target information) to the mKFT with different sets of EMGs, in an attempt to simulate the signals that might be available at different levels of spinal cord injury (SCI). To simulate a C4 injury we used just the upper trapezius, we added the deltoids, brachioradialis, biceps and pectoralis major for C5, and for C6 we also added the triceps [12]. Algorithm accuracy was quantified by the multiple R^2 [13].

C. *Sensitivity to Errors in Target Estimates*

Due to attentive subjects and idealized experimental conditions, the gaze recordings in our experiments were of reasonably high quality. This is unlikely to be true outside of a laboratory setting. We therefore simulated target estimates of varying quality to evaluate the effect on performance. This was done by assigning two potential targets to each of the reaches: its correct target and another which was randomly selected from the set of all targets. We evaluated the accuracy of the mKFT where the prior probabilities assigned to the correct and incorrect targets were varied, again at different levels of simulated SCI.

D. *Adding a Generic Mixture Component*

In the worst-case-scenario, which would occur if the subject looked in a completely different location, the correct target would have a zero prior. In this case, with no knowledge of the target, our best possible estimate of the state would come from the generic KF. To take advantage of this, we added a generic mixture component with a prior probability of 0.1; either the correct or random targets were given a prior of 0.9, thereby simulating the extremes of the analysis described above. If the EMG signals provide strong evidence that the generic KF is more appropriate than the KFT component, the weight for the generic component will converge to 1 through the likelihood term. Our solution will thus be close to that of the generic KF, which is preferable when the KFT has been initialized with an inaccurate target.

III. Results

A. Decoding Performance with Gaze and EMG

For all simulated SCI levels, the mKFT outperformed the KF. The accuracy of the mKFT was quite consistent from C6 - C5, however it was reduced somewhat for C4. The generic KF degraded more dramatically at the simulated C4 level, as would be expected from a non-directional model with only a single EMG source (Fig 2A).

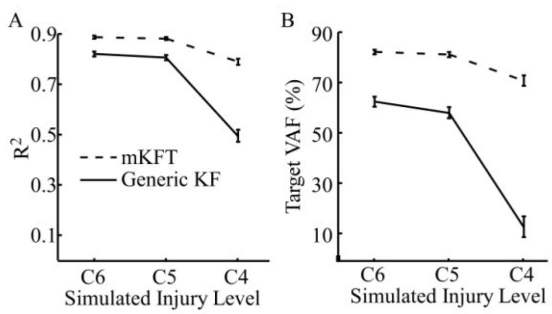

Fig. 2. Mean and standard error for A Trajectory R^2 and B Target VAF for mKFT and generic KF with different simulated injury levels, using gaze information for target priors.

Because all targets were in front of the subject, a substantial component of the R^2 was related to the outward component of the reach common to all targets. This can be seen in the R^2 of nearly 0.5 for the generic KF at C4; the decoded trajectories for this condition had errors similar to those for a decoder that consistently predicted a reach to an average target location, regardless of the actual target. To better evaluate final reach accuracy we also calculated the target variance accounted for (VAF) by scaling the squared error at the end of the reach by the variance of the target locations. Using this measure (Fig 2B), it can be seen that the generic KF at C4 is unable to decode endpoint target location. This is illustrated in the example reach shown above (Fig 3), where the generic KF at C4 does well in the vertical (Z) and outward (X) directions, but its lateral prediction (Y) is completely inaccurate. Bearing this in mind, we will report multiple R^2 for the trajectory for the remainder of the paper to provide a measure of decoding accuracy encapsulating all dimensions relevant to our experiment.

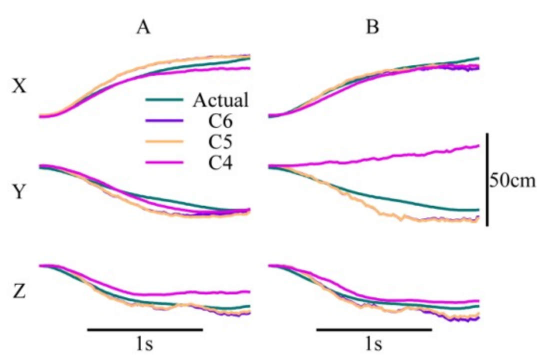

Fig 3. Sample reach and predictions of A mKFT and B generic KF for different simulated injury levels.

B. Simulated Target Estimates

Unsurprisingly, performance of the mKFT was best when the correct target was given a prior probability of 1 (the KFT case). There was a small drop in accuracy when the random target was given a nonzero prior. As the prior for the correct target was reduced from 0.9 to 0.1, the drop in accuracy was relatively small in the simulated C6 - C5 cases. However, at C4 the influence of the assigned priors became more evident (Fig 4). With only the upper trapezuis EMG with which to perform the likelihood calculation, the decoder does not converge to the correct trajectory as quickly or reliably. Nonetheless, the mKFT was significantly more accurate than the generic KF so long as the correct target was represented in the mixture, no matter how small its prior. However, when the random target was given a prior of 1, the KFT performance dropped well below that of the generic KF.

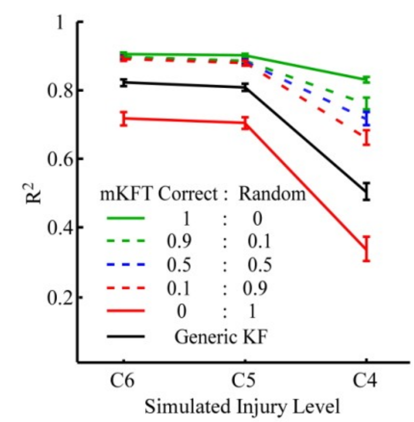

Fig. 4. Mean and standard error R^2 for the generic KF and mKFT, for which the prior probabilities assigned to the correct and random targets are shown in the legend.

C. Adding the Generic Mixture Component

Adding the generic KF into the mixture greatly reduced the errors of the mKFT with the incorrect target (Fig 5A). At the simulated C6 and C5 levels, it performed as well as the generic KF, and at C4 it was only slightly worse. While the weight for the generic KF was initialized at 0.1, it quickly increased as likelihood of the EMG provided evidence that the KFT component was less accurate (Fig 5C). There was, however, a small cost to the accuracy of the mKFT where the target was correct, as a higher weight was assigned to the generic KF for a low proportion of the time (Fig 5B).

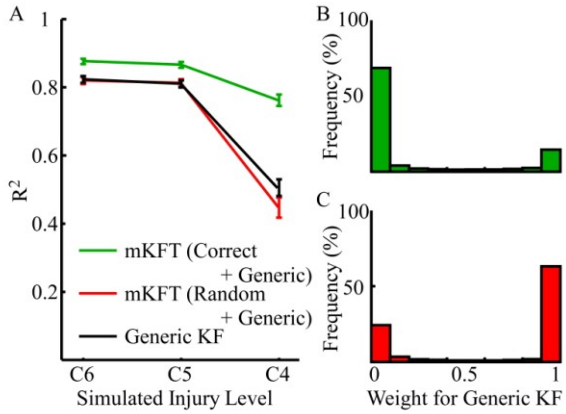

Fig. 5. A: Mean and standard error R^2 for generic model and mixture of generic model and KFT with correct and random targets. B, C: Histogram of the weights assigned to the generic model in the C6 case, for the correct and random target mixtures respectively.

IV. Discussion

Incorporating eye tracking information significantly improved our decoding performance, particularly when there was very limited EMG available. When we don't have many neural signals to decode from, a strong choice of trajectory model becomes even more important. We cannot expect a patient's gaze information to be highly informative all the time, however. Here we simulated target information of varying quality to see how our approach might perform in a more realistic functional environment. We found that the mixture model performed well when faced with noisy target information. By incorporating an additional generic KF into the mixture, we avoided a degradation in performance when the target estimate was incorrect.

The increased sensitivity of our mixture model to the target priors when the EMG is limited may present a problem for implementing our approach in the high tetraplegia population, where there are few remaining control signals. Over-reliance on the eye movement recordings would be unacceptable, and it is critical that our proposed system does not impinge on the functionality that remains available to the patient population. Encouragingly, there was a significant benefit to incorporating the target information even when the prior for the correct target was very low.

In a closed-loop BMI application, the user can modify their behavior to improve decoding as they receive feedback. We expect that the inclusion of target information will reduce their cognitive burden by increasing the decoder's reliance on the trajectory model relative to the continuous control signals. If the target estimates are all inaccurate however, it is important that we avoid making inappropriate predictions that worsen the situation. We think that by including the generic KF in the mixture we could provide the user with the opportunity to overcome these effects by relying more heavily on their voluntary control signals.

Acknowledgment

The authors thank T. Haswell, E. Krepkovich, and Dr. V. Ravichandran for their assistance with experiments.

References

[1] R. S. Johansson, G. Westling, A. Bäckström, and J. R. Flanagan, “Eye–hand coordination in object manipulation,” *the Journal of Neuroscience*, vol. 21, no. 17, p. 6917, 2001.

[2] R. J. K. Jacob and K. S. Karn, “Eye tracking in human–computer interaction and usability research: ready to deliver the promises (section commentary),” in J. Hyona, R. Radach, & H. Deubel (Eds.), *The Mind’s Eyes: Cognitive and Applied Aspects of Eye Movements.*, Oxford: Elsevier Science, 2003, pp. 573-605.

[3] E. A. Corbett, E. J. Perreault, and K. P. Körding, “Mixture of time-warped trajectory models for movement decoding,” in *Advances in Neural Information Processing Systems*, Vancouver, 2010, vol. 23.

[4] B. M. Yu et al., “Mixture of trajectory models for neural decoding of goal-directed movements,” *Journal of neurophysiology*, vol. 97, no. 5, p. 3763, 2007.

[5] 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, p. 2465–2494, 2006.

[6] C. Kemere and T. Meng, “Optimal estimation of feed-forward-controlled linear systems,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2005, vol. 5, pp. 353-356.

[7] G. H. Mulliken, S. Musallam, and R. A. Andersen, “Decoding Trajectories from Posterior Parietal Cortex Ensembles,” *the Journal of Neuroscience*, vol. 28, no. 48, pp. 12913-12926, 2008.

[8] G. Wu et al., “ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion–Part II: shoulder, elbow, wrist and hand,” *Journal of Biomechanics*, vol. 38, no. 5, p. 981–992, 2005.

[9] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Journal of basic Engineering*, vol. 82, no. 1, p. 35–45, 1960.

[10] W. Wu, Y. Gao, E. Bienenstock, J. P. Donoghue, and M. J. Black, “Bayesian population decoding of motor cortical activity using a Kalman filter,” *Neural Computation*, vol. 18, no. 1, p. 80–118, 2006.

[11] J. E. Kulkarni and L. Paninski, “State-space decoding of goal-directed movements,” *Signal Processing Magazine, IEEE*, vol. 25, no. 1, p. 78–86, 2007.

[12] F. M. Maynard et al., “International standards for neurological and functional classification of spinal cord injury,” *Spinal cord*, vol. 35, no. 5, p. 266–274, 1997.

[13] E. J. Perreault, R. F. Kirsch, and A. M. Acosta, “Multiple-input, multiple-output system identification for characterization of limb stiffness dynamics,” *Biological cybernetics*, vol. 80, no. 5, p. 327–337, 1999.