Generative Decoding of Intracortical Neuronal Signals for Online Control of Robotic Arm to Intercept Moving Objects

Chenyang Li, Yiheng Zhang, Tianwei Wang, Xinxiu Xu, Qifan Wang, Bradley Xu and He Cui
CAS Center for Excellence in Brain Science and Intelligence Technology and Institute of Neuroscience, University of Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai, 200031 P.R. China
cuihe@ion.ac.cn

Abstract. Brain-machine interfaces (BMIs) are being developed for translating neural signals to control an external device such as a robotic arm. The prevalent decoding strategy of current BMIs is the continuously conversion of neural spike trains to motor variables via a discriminative algorithm, but this typically results in unnatural and slow robotic motion. Recently, we implemented a BMI under feedforward control for interception of moving targets, though the decoding accuracy was insufficient for accurate online control. To improve the BMI performance, a generative algorithm called Latent Factor Analysis via Dynamical Systems (LFADS) [1], was applied to data pre-processing for de-noising. The results indicate potential advantages of feedforward control with generative decoding in improving BMI design.

1. Introduction

Over the last two decades, invasive BMI emerged as a promising approach to neuroprosthetics for patients with brain disorders affecting motor control by linking neural activity to external actuators, such as a robotic arm. By continuously converting population neuronal activity recorded from multi-channel electrode arrays to physical movement parameters with discriminative decoders, most BMIs to drive prosthetic devices rely on feedback control without prior trajectory planning [2–4]. Consequently, their performance in motor speed, flexibility, and smoothness falls short of clinical feasibility.

As a valuable animal model for the neural control of movement and BMI, the monkey is able to generate appropriate motor commands driving sophisticated skeletomuscular system in response to signals from the complex and dynamic environment. Recently, we demonstrated that monkeys can utilize a feedforward control strategy during manual interception of moving targets [5], and have implemented a feedforward BMI that interacts with the dynamic environment [6].

In contrast to the traditional point of view, accumulating evidence has support the dynamical perspective that complicated movement parameters are generated by the dynamical evolution of initial premovement neural states [7]. In principle, the motor cortex generates motor commands rather than simply describing motor variables. Therefore, decoding movement preparatory activity with a generative algorithm for feedforward control seems a feasible biomimetic design for clinically viable BMI.

In the present study, we have added a generative model, LFADS proposed by Pandarinath et al. [8], into neural decoding. This encoding-decoding framework is able to generate population neural data, and its encoder/decoder can capture neural dynamics. Based on this model, potential population neural data
can be estimated as data pre-processing, so that decoded signals can accurately control a robotic arm to rapidly catch a moving object in real-time.

2. Method

2.1. Behavioral paradigm

Two male rhesus monkeys (Macaca mulatta, 5 and 8 kg) were trained to perform the flexible interception task shown in Figure 1. First, a target (a green dot) appeared at the centre of the screen and the monkey was required to hold its hand on it. After 600 ms, a green dot appeared at a random position on a 9-cm radius circle, moving in one of 5 randomly-chosen angular velocities (-240 degree/s, -120 degree/s, 0 degree/s, 120 degree/s, 240 degree/s). During that period, the monkeys were required to hold the central target for 400 to 800 ms. Then, the central target turned dark as a ‘GO’ signal to allow the monkeys to make an interception. If the error between hand position and moving dot was less than 2.4 cm, a juice reward was delivered. In each session, the monkeys were guided to sit in a primate chair positioned 30 cm in front of a touch-sensitive screen (Elo). The height of the screen centre was adjusted to the monkey’s shoulder.

![Figure 1](image)

**Figure 1.** Sketch map of a monkey performing interception task. The monkey first held its hand on a central dot for 600 ms. Then, a dot appeared at a random position on a circle moving clockwise (CW) or counter-clockwise (CCW) at one of 5 speeds. The central dot turned dark as a ‘GO’ signal to initiate the interception. Modified from Li et al. 2018 [2]

2.2. Surgical procedures and recording system

Before recording sessions, two 96-ch microelectrode arrays (Blackrock Microsystems, Salt Lake City, UT) were implanted into the motor cortex and posterior parietal cortex (Figure 2.). The M1 array is located in the anterior bank of the central sulcus, lateral to the precentral dimple. The other was implanted between the postcentral dimple and medial bank of the intraparietal sulcus (IPS). The population neuronal activity was captured using Blackrock Microsystems, sampling at 30 kHz. A band-pass filter (0.3 to 9 kHz) was used to filter each channel’s signal, and online spike sorting was set manually. All surgical and experimental procedures were approved by the Institutional Animal Care and Use Committee.
2.3. Online decoding system

An integrative system was built to collect and decode population neuronal activity, EMG, and hand motion. Recorded signals by Blackrock Microsystems were extracted by the Python module, called cbpy. EMG signals were recorded with the Trigno Avanti Platform (Delsys, Natick, Massachusetts). Arm movement signals were captured by the Motion Capture System (Vicon, Centennial).

As shown in figure 3, the output from the decoder was sent to an actuator, a robotic arm (Universal Robots). A spin wheel was positioned in front of the arm, and 12 LEDs were uniformly placed on its circumference to simulate the moving dots that appeared when the task was performed on the touch screen. Through TCP/IP, the decoded signal was sent to the robotic arm and hand, driving them to execute the decoded movement to reach and grasp the moving LED. The monkey’s performance was compared to that of the robotic hand to test the feasibility of this BMI.

![Figure 2. Implanted locations of micro-electrode arrays. Two 96-ch microelectrode arrays were implanted in M1 and PPC. Posterior central dimple and intraparietal sulcus restricted the implant site in PPC as area 5d. Meanwhile, the central sulcus and prior central dimple led to the arm area for the other array.](image)

![Figure 3. Integrative online decoding system. All equipment communicated with each other via a switch. This system contained multiple signal collector for real-time capture of neural signal, movement parameters and EMG signal. The controller PC and the decoding PC send the control signal to spin wheel and robotic arm respectively to mirror BCI performance.](image)
2.4. The design of decoder

In contrast to most BMIs, a generative model under the feedforward control framework was utilized to enhance decoding accuracy. Neural activity within a time window of 500 ms before movement onset was extracted and converted to spike counts in 10 time-bins to preserve temporal features. A Wiener filter and support vector machine (SVM) were adopted to test the performance of the BMI.

The Wiener filter was an optimal linear filter that could be used for linear translation of spike counts to movement parameters. The Finite Impulse Response (FIR) Wiener filter as an optimal filter could be expressed as:

\[
s(t) = b + \sum_{i=0}^{n} h(u\Delta)x(t - u\Delta),
\]

where \( s \) is the movement parameters. In this study, \( s \) was the endpoint position touched by the monkey, \( h \) the impulse response function, \( b \) the intercept, \( x \) the spike count of all recording channels, and \( u\Delta \) the \( u \)-th time interval. The least mean square (LMS) method was used to get model parameters. To avoid overfitting, L1-regularization was applied to model training and the regularization term was set by grid-search.

Another strategy to convert neural signals to movement parameters was to design a linear classifier. Therefore, C-SVM with a linear kernel was a reasonable choice to be applied to decoder design [9]. One SVM classifier could classify two clusters, and its expression was

\[
f(x) = sgn(w^Tx + b),
\]

where \( x \) was the same decoder input as the Wiener filter, and \( w \) and \( b \) were parameters of the linear kernel. The classification was achieved by the sign function. To avoid over-fitting, we used grid-search to set proper soft margin parameters.

For training classifier, the monkey’s touched positions were divided into 18 classes, the difference between two classes being 20 degrees. The classification of multi-classes used a one-versus-one strategy.

As a common linear regression model, the Wiener filter used LMS to fit neural activity and position touched by the monkey. The solution was the product of the auto-correlation function of input and cross-correlation function of input and touched position. Because of the fixed time window for decoding, application of the Wiener filter was appropriate in our feedforward control framework. The goal of the SVM classifier is to define a hyperplane to achieve classification. If the neural activity could be split into several clusters, SVM would work well, although a continuous distribution of neural activity in high dimensions would weaken the classification ability of SVM. Another advantage of SVM is robustness, since a well-trained hyperplane can be used to infer the cluster of underlying neural activity, or in recognizing possible noise. These characteristics might benefit the achievement of volitional control, and its many alternative nonlinear kernel functions would give help find nonlinear relationships between the input and output.

LFADS [1] is a variational autoencoder using a recurrent neural network as encoder and decoder. This model contributed to the estimation of potential neural dynamics, and its framework helped to infer more potential neural data. Because of its denoising characteristics, this model can improve BMI performance by reducing trial-to-trial variability.

Given these considerations, the decoder framework was designed as Figure 4. First, 400 trials were collected for LFADS training. Meanwhile, an extra 300 trials were collected. Then, these two datasets were combined as a training set, and duplicated as two sets for LFADS pre-processing and smoothing with a Gaussian kernel. Finally, the two pre-processed datasets were concatenated and fed to last stage of the decoder.
Figure 4. Framework of generative decoding and feedforward control. Population neural data within 500 ms before movement onset was used to train a generative (LFADS) model and a discriminant model (Wiener filter). The training sets were duplicate, and then pre-processing by two algorithms: data reconstruction, and data smoothing. The pre-processed datasets would concatenate to one set as input to the Wiener filter.

3. Results
Compared to previous results, the BMI decoding accuracy was significantly increased. Nearly 70 recording channels were chosen for feeding the decoder. The standard to deciding if the decoding result was successful was that the difference between the decoding movement direction and the actual movement direction should be less than 30 degrees. In our previous study [6], the SVM decoder performed better than the Wiener filter, but here, the Wiener filter showed better performance. Therefore, we compared the best decoder in each year as shown in Figure 5. Overall, decoder performance significantly increased, from 66.2% to 87.1%.

As shown in Figure 6, the Wiener filter with L1 regularization (LASSO) overperformed SVM. Different colours in Figure 5. indicate the related data pre-processing methods: blue the concatenated dataset of data pre-processed by LFADS and convolved by the Gaussian kernel, orange the dataset that is only pre-processed by LFADS; and yellow another pre-processing method. In the SVM group, LFADS contributed to improving BMI performance. However, in the Wiener filter group, these three data pre-processing methods showed similar decoding performance, while the concatenated data performed best with a small gap. Therefore, the neural signal should be the main reason for the improvement, and using concatenated data for decoding became the most appropriate pre-processing method.
Figure 5. Increase in decoding accuracy from 2018 to 2019. Chance level was 16.67%.

Figure 6. Histogram of decoding performance of two decoders with different data pre-processing methods. Using the same dataset, the success rate of the Wiener filter was significantly higher than SVM, and the concatenated data (blue) gave the best decoding performance. Chance level was 5.56%.

4. Discussion

Feedforward control provides a more plausible framework to implement LFADS in online decoding. As a powerful generative model, LFADS is based on the assumption that spike counts in every time interval can be reconstructed by a well-trained recurrent neural network and a Gaussian distribution (the latent factor). This design might limit its application in asynchronous BMIs under feedback control, because the Gaussian distributions often are different in each time bin, whereas the model parameters are fixed. Although the current improvement derived from LFADS was modest, it indicates the potential feasibility of LFADS in our synchronous BMI under feedforward control.

Comparing SVM to the Wiener filter indicates that linear classification was not appropriate to BMI decoding. High-dimensional vector input could be a reason for this advantage. Using PCA for decomposition could increase the decoding success rate, but still was not the best decoding strategy, implying that linear or nonlinear dimensionality reduction is an alternative method for decoding improvement. In particular, using grid-search to compare the success rate between different kernel functions, the polynomial kernel performed better. Therefore, there was a nonlinear relationship between populational activity and movement parameters. Combined with the generative model, mapping motor cortex neural activity to a manifold for decoding is another possible method.

To optimize BMI performance, there are still some potential improvements. First, LFADS can infer many underlying population activities, but there is still demand for a proper decoder to recognize the mixed recorded data and inferred data to make use of the neural data. Second, this BMI was a synchronous BMI, which relied on an external trigger. Thus, the detection of movement intention was the subsequent task to achieve. Third, in order to accurately intercept the LEDs on the spin wheel, the robotic arm should have been able to immediately respond to control instructions.

5. Conclusions

Due to of more neurons and better signal quality, decoding accuracy significantly increased 66.2% to 87.1% in the present study. Furthermore, the feasibility is demonstrated of applying the generative model, LFADS, to online decoding in synchronous BMIs. Therefore, generative decoding under feedforward control appears to be a promising approach for accurate and flexible BMI control.
especially for interactions with dynamic environments.

6. Acknowledgments
The authors would like to thank Chen Zhao, Chao Guan and Peiwen Ding for their valuable advice and comments. This work is supported by National Key R&D Program of China, (Grant No. 2017YFA0701102) and Shanghai Municipal Science and Technology (Grant No. 18JC1415100).

References
[1] Pandarinath C, O'Shea DJ, Collins J, et al. Inferring single-trial neural population dynamics using sequential auto-encoders. *Nat Methods* 2018; 15: 805–815.
[2] Costa RM, Ganguly K, Costa RM, et al. Emergence of Coordinated Neural Dynamics Underlies Neuroprosthetic Learning and Skillful Control. *Neuron*. Epub ahead of print 2017. DOI: 10.1016/j.neuron.2017.01.016.
[3] Gilja V, Nuyujukian P, Chestek CA, et al. A high-performance neural prosthesis enabled by control algorithm design. *Nat Neurosci*. Epub ahead of print 2012. DOI: 10.1038/nn.3265.
[4] Velliste M, Perel S, Spalding MC, et al. Cortical control of a prosthetic arm for self-feeding. *Nature* 2008; 453: 1098–1101.
[5] Li Y, Wang Y, Cui H. Eye-hand coordination during flexible manual interception of an abruptly appearing, moving target. *J Neurophysiol* 2018; 119: 221–234.
[6] Li C, Zhang Y, Wang T, et al. An Implementation of Brain-Machine Interface by Decoding Predictive Intracortical Signals toward a Moving Object *. 2018 IEEE Int Conf Cyborg Bionic Syst* 2018; 520–523.
[7] Churchland MM, Cunningham JP, Kaufman MT, et al. Neural population dynamics during reaching. *Nature*. Epub ahead of print 2012. DOI: 10.1038/nature11129.
[8] Pandarinath C, O'Shea DJ, Collins J, et al. Inferring single-trial neural population dynamics using sequential auto-encoders. *Nat Methods* 2018; 15: 805–815.
[9] Fan R, Chen P, Lin C. Working Set Selection Using Second Order Information for Training Support Vector Machines. 2005; 6: 1889–1918.