Feasibility Analysis of Genetically-Encoded Calcium Indicators as a Neural Signal Source for All-Optical Brain-Machine Interfaces

Xulu Sun, Jonathan C. Kao, IEEE Student Member, James H. Marshel, Stephen I. Ryu, IEEE Member, and Krishna V. Shenoy, IEEE Senior Member

Abstract— Optical techniques such as two-photon (2p) calcium imaging have the potential to transform the way we interrogate neural circuits, both in the realm of basic neuroscience and in the development of brain-machine interfaces (BMIs). This may be possible by overcoming some of the limitations of electrophysiological methods. Here we ask if optical imaging signals, in particular 2p GCaMP6 calcium imaging signals, can benefit BMIs despite their relatively long activity-response time constants, low signal-to-noise ratios (SNRs), and slow acquisition frame rates. We employed motor cortical electrode array recordings as the basis for generating synthetic 2p GCaMP signals. We then decoded movement kinematics from these surrogate data using a state-of-the-art BMI decoder algorithm. We found that it was possible to decode the position and velocity of the hand from synthetic imaging signals. We quantified the decoder performance using standard mean squared errors (MSEs) and Pearson’s correlation coefficient (r) measures. Decode quality varied considerably as a function of SNR and the frame rate of data acquisition. Future computational and experimental research is required to quantify SNR more accurately and to increase the imaging frame rate while maintaining high SNR, in order to improve all-optical BMI (o-BMI) performance. This study should help establish the feasibility and design space of o-BMIS.

I. INTRODUCTION

BMIs aim to restore lost motor function for people with neurological injury or disease. To date the best-performing intra-cortical BMIs have decoded movement intention from neuron action potentials (spikes) measured with multi-electrode arrays. Such BMIs have enabled clinical trial participants to control computer cursors [1], [2] or robotic arms [3], [4]. Despite this recent progress, performance is inherently limited by being able to record electrically from only hundreds of neurons simultaneously using electrodes spaced hundreds of microns apart. In order to restore near-natural levels of motor function, many more simultaneously recorded neurons are needed (e.g., many thousands [5]). Further, the ability to reliably record from the same neurons across weeks to years should allow for BMI robustness across time as well as neural adaptation [6].

To explore the possibility of overcoming this limit, it is timely to consider the feasibility of an o-BMI in nonhuman primates (NHPs) which are the central animal model for BMIs [7], [8] following what was recently explored in mice [9]. An o-BMI provides the potential to record and decode population responses from many thousands of neurons stably and simultaneously. This larger pool of neurons would enable the study of how BMI performance scales with the number of neurons available to the decoder. O-BMIS also provide a means to test how performance varies with specific neuron cell types and in dense populations of neurons of local circuits. These comparisons enabled by optical imaging should give rise to a wealth of opportunities for algorithmic and signal processing improvement to facilitate better BMI performance [5].

Given the promising advantages of optical methods over traditional electrophysiology, there have been increasing reports of translating optical imaging techniques to NHPs including marmosets [10] and rhesus macaques [11]. In particular, rhesus macaques are the principal animal model for BMIs due to their human-like behavioral repertoire and neuroanatomy. While systems for in vivo optical imaging in the motor cortex of behaving monkeys are progressing, they have not yet been fully developed [12]. As a result, a rigorous quantitative comparison between imaging and electrode data ideally from the same animal, as well as their BMI decoding performance (both offline/open-loop and online/closed-loop) has yet to be performed.

The present study aims to provide a general framework, using available electrophysiological data and estimates in conjunction with initial simulations, for considering the challenges and opportunities in designing of o-BMIS. We do so by decoding hand trajectories from surrogate calcium imaging signals that are generated from multi-electrode array recordings. By using trajectory reconstruction performance as a measure of kinematic information in the neural signals, we evaluated how well surrogate 2p calcium imaging signals could preserve kinematic information encoded by the original neural spikes. We found that, under appropriate sampling rates and noise levels, decoding surrogate imaging signals to generate hand trajectories was possible despite calcium indicators having (1) relatively slow response kinetics to action potentials, (2) various sources of noise, and (3) non-linearity, including saturation, between the neural spikes and...
in intracellular calcium concentration. We also document that performance of the simulated o-BMIs was sensitive to frame rate of data acquisition (e.g., hundreds of frames/s understandably provided better decoding than tens of frames/s). These results provide insights into the potentially promising use of o-BMIs in NHPs and, more broadly, into the value of applying optical techniques in NHP studies.

II. METHODS

A. Behavioral Tasks

All procedures and experiments were approved by the Stanford University Institutional Animal Care and Use Committee. An adult male rhesus macaque (J) was implanted with two 96 electrode Utah arrays (Blackrock Microsystems Inc., Salt Lake City, UT) using standard neurosurgical techniques, one in prefrontal cortex (PMd) and the other in primary motor cortex (M1). Monkey J was trained to make point-to-point reaches on a center-out-and-back task in a 2D plane with the arm contralateral to array implants. The behavioral paradigm was the same as recently reported [13] where the monkey reached from a center target to one of eight peripheral targets and reached back to the center. Eight peripheral targets were equally spaced along a 12 cm radius circle and only one was prompted on each single trial. Hand position data were measured with an infrared reflective bead tracking system (Polaris, Northern Digital, Ontario, Canada).

B. Physiological Models for Calcium Indicators

A simple biophysical model for calcium imaging data assumes that the intracellular calcium concentration rises instantaneously after one action potential, and decays exponentially with a time constant \( \tau \) determined by the intrinsic property of calcium indicators (Fig. 1). It also assumes that each jump of calcium concentration caused by each action potential is of the same amplitude, and the fluorescence signals emitted by calcium indicators change linearly with intracellular calcium concentration. Furthermore, previous work suggests that systematic noise in imaging data follows a Gaussian distribution [14]. We now briefly present the key governing equations.

Assume that the baseline (resting) calcium concentration in each neuron is a constant \( [\text{Ca}^{2+}]_{\text{rest}} \) nM and that \( [\text{Ca}^{2+}]_\text{t} \) the intracellular calcium concentration at time \( t \) increases by \( \Delta \text{Na} \) nM after each spike and then exponentially decays with time constant \( \tau \) ms. Given these assumptions, the calcium concentration dynamics are given by:

\[
[\text{Ca}^{2+}]_{t+1} = (1 - \Delta/t)[\text{Ca}^{2+}]_t + (\Delta/\tau) [\text{Ca}^{2+}]_{\text{rest}} + A_n_t
\]

where \( \Delta \) is the time step size (i.e., the data acquisition frame period), and \( n_t \) is the number of spikes having occurred between time \( t \) and \( t+1 \). Assuming that fluorescence is linearly related to \( [\text{Ca}^{2+}] \) in the absence of noise and under the linearity approximation, at time \( t \) the corresponding fluorescence measurement \( F_t \), which is the measured calcium signal, can be written as

\[
F_t = \alpha[\text{Ca}^{2+}]_t + \beta + \epsilon_t \sim N(0, \sigma^2)
\]

where \( \alpha \) indicates a linear scaling from calcium concentration to fluorescence signal, \( \beta \) represents baseline fluorescence level, and \( \epsilon_t \) is the noise at time \( t \) and follows a time-invariant Gaussian distribution that is assumed to be independent and identically distributed (i.i.d.) through time. Note that \( A \) and \( \alpha \) can be absorbed by one constant scalar indicating the change of \( \Delta F/F \) caused by one action potential, indicated below as \( (\Delta F/F)_{\text{unitary}} \); \( [\text{Ca}^{2+}]_{\text{rest}} \) and \( \beta \) can be absorbed by one constant indicating the background

Figure 1. An overview of the procedure for calcium imaging data simulation. The spike train of one example neuron (A, first panel) was convolved with a single or double exponential decay kernel (B), which is a common model for spike-calcium conversion. Gaussian noise at different levels (C) was then added to the convolved signal (A, second panel) to generate the surrogate fluorescence signal (A, third panel). The fluorescence signal was then down-sampled at different rates to simulate the lower sampling rate of imaging techniques compared to electrode recordings (A, fourth panel). Fluorescence signal saturation, if considered, was computed before adding Gaussian noise.
fluorescence signal. Empirical $[\text{Ca}^{2+}]_{\text{rest}}$ is around 50 nM [15], the value we used in this study.

This simplified model does not deal with the non-linearity of calcium indicators, especially fluorescence saturation that could be a common source of signal corruption when the spike rate is high (e.g., an activated neuron in rhesus M1 can exhibit firing rate over 100 spikes/s, while mouse neuron firing rates tend to be lower). To deal with the non-linearity between calcium concentration and fluorescence signal, we replaced the linear equation (2) by the standard single wavelength conversion equation that is typically used to fit the calcium-fluorescence relationship of calcium dyes like OGB-1[16, 17]:

$$[\text{Ca}^{2+}] = \frac{[\text{Ca}^{2+}]_{\text{rest}} + K_d (\Delta F/F)_{\text{max}}}{1 - (\Delta F/F)_{\text{max}}}$$ (3)

where $[\text{Ca}^{2+}]_{\text{rest}}$ is the same as previously described, $\Delta F/F$ is the relative change of fluorescence intensity normalized to the baseline fluorescence level, $(\Delta F/F)_{\text{max}}$ is the maximum range of fluorescence change of the calcium indicator, and $K_d$ is the dissociation constant of the calcium indicator. Equation (3) can be re-written as:

$$\Delta F/F = (\Delta F/F)_{\text{max}} \frac{([\text{Ca}^{2+}]_{\text{rest}} + [\text{Ca}^{2+}])}{(1 + [\text{Ca}^{2+}]_{\text{rest}} + K_d)}.$$ (4)

### C. Simulation of Calcium Imaging Data

**Convolution of original spikes.** Without noise added, the conversion from spikes to calcium concentration (and to fluorescence signal), as suggested by the simplest model in equations (1) and (2), is an exponential decay convolution of the original spike train data:

$$f(t) = A e^{-n t_{\text{off}}/t}$$ for $t \geq t_0$

$$f(t) = 0$$ for $t < t_0$$ (5)

where $A$ is increase in $\Delta F/F$ caused by one action potential, $t_0$ is the time when a spike occurs, and $t_{\text{off}}$ is the time constant for exponential decay (Fig. 1B).

When the finite rise time due to slower calcium binding to indicators was taken into account, we used the following convolution kernel:

$$f(t) = A \left(1 - e^{-t_{\text{off}}/t_0}\right) e^{t_{\text{off}}/t_0}$$ for $t \geq t_0$

$$f(t) = 0$$ for $t < t_0$$ (6)

where $t_{\text{on}}$ is the rise time constant for a certain calcium indicator. Typical rise times for calcium dyes like OGB-1 are around 10 ms [17], and the rise times for commonly used genetically-encoded calcium indicators (GECIs) measured in mouse V1 in vivo have recently been reported (e.g., 45 ms for GCaMP6f, used in our work for the convolution kernel) [18].

**Saturation.** Saturation equations differ depending on the kinetics of calcium indicators. Here we only considered the widely-used and well-characterized saturation equation for calcium dyes as indicated by equation (4). When saturation was incorporated in the simulation, convolution was done for conversion from spikes to calcium concentrations as described above, and the convolved calcium concentration served as $[\text{Ca}^{2+}]$, in equation (4) for the computation of $\Delta F/F$. Because $(\Delta F/F)_{\text{max}}$ is a constant scalar and should not change the decoder’s performance, we arbitrarily set it to 1 in the simulation without loss of generality. Note, we did not explicitly investigate the potential effects of non-linearity, although our addition of saturation should simulate the decreased jump size of $\Delta F/F$ of one action potential as firing rates increase, and the addition of Gaussian noise described below should simulate the effect of low SNR on not detecting low firing rates in the optical signal.

**Noise.** There are many sources of noise in real 2p imaging data, including the fluctuation of baseline calcium concentration, the dynamics of calcium-GCaMP binding, the number of released photons that follow a Poisson process, the optical collection and measurement noise in microscope imaging systems, etc. (see Table 1). Here we grouped all sources together as a single Gaussian noise term with zero mean and variance $\sigma^2$. Lacking firm estimates of most of these constituent noise sources, a simple and appropriate estimate to include at this early stage of investigation is a single Gaussian noise source. This is motivated by the central limit theorem and supported by recent results [14]. We added Gaussian noise to the convolved signal when...
saturation was not included (i.e. the convolution converted spikes linearly to ΔF/F), or to the saturated signal computed by equation (4). We also varied noise level by varying the ratio $\sigma/(\Delta F/F_{\text{unitary}})$ where $(\Delta F/F)_{\text{unitary}}$ is the unitary peak change of $\Delta F/F$ caused by one action potential.

**Down-sampling.** The simulated data generated in the first three steps as described above had the same sampling rate (1000 Hz) as the original spike train data. To simulate the temporal resolution of 2p imaging data, we then down-sampled at 200, 100, or 30 Hz, which are achievable 2p imaging rates in mice [17], [19]. Both the non-down-sampled 1000 Hz data and down-sampled data were passed to the BMI decoder for hand trajectory estimation (reconstruction) and performance quantification.

**D. Neural Decoding and Performance Quantification**

All measures of the decoder performance in this study come from decodes of previously collected offline electrophysiological data from native arm behavioral trials rather than from closed-loop online BMI experiments where a neurally-controlled cursor would be used to move and touch a visual target. We decoded both spike train data and surrogate calcium data in order to conduct a head-to-head comparison. We also removed different numbers of electrodes (10-190 channels “dropped”) in reverse order of importance defined by mutual information between the spike distributions and reach direction. We then decoded hand kinematics at different electrode counts, as shown in Fig. 2A. Spike train data were recorded while monkey J performed native arm movement tasks as described in section II.A. Note that a -4.5 RMS threshold was used to detect spiking activity on an electrode, as opposed to spike sorting to identify action potentials with individual neurons. Surrogate calcium imaging data were generated from the original spike data as described in sections II.B and II.C. We used the Feedback Intention Trained Kalman Filter (FIT-KF) algorithm [20] to decode all data. Briefly, FIT-KF augments a standard velocity Kalman filter by incorporating assumptions about the nature of closed-loop neural prosthetic control and by adjusting kinematics of the training data to better match movement intention [7], [20]. We did not incorporate the position correction step of the FIT-KF. We binned neural data from the previous 15 ms and output a velocity command at the end of every binning window. We trained the decoder with 80% of the 500 trials recorded within one session, and tested its performance with the remaining 20% of trials. The decoder performance was measured by mean squared error (MSE) between decoded hand trajectories and real hand trajectories, and by Pearson’s correlation coefficient ($r$) of decoded hand velocity and real hand velocity. Both measures were averaged across all test trials.

**III. RESULTS**

**A. Overall Comparison of Decoding Spike Data and Surrogate Imaging Data**

As expected, decoding performance for both spike train data and surrogate 2p imaging data improved as the number of recorded neurons increased (Fig. 2A). Decoded imaging data (200 Hz sampling rate) did not appear to perform (larger MSE and smaller $r$) as well as decoded spikes. Importantly, the decoded hand trajectory was reasonably accurate based on previous work that reported that a Pearson’s $r$ above 0.7 typically indicates reasonably good online decoding performance [21]. We achieved $r=0.68$ from decoding synthetic $\Delta F/F$ when the imaging data were generated with a noise level of $\sigma/(\Delta F/F)_{\text{unitary}}$ equal to 1 and a 200 Hz sampling rate (see Fig. 2B, example reconstructions of an “upper-left” reach trial and a “right” reach trial). These noise levels and sampling rates are achievable with current imaging techniques though with a small field of view [19]. We emphasize that although the offline decoding performance is reasonable, experiments are needed to assess the closed-loop performance of the system. Given that advanced imaging techniques could enable recording from larger numbers of neurons, we also envision that o-BMI decoding performance could continue to increase beyond what is currently possible with electrical recordings (though, importantly, electrical recordings are also scaling up). An important caveat is that if trends shown in Fig. 2A are sufficiently accurate (e.g., decoder performance saturation for a 2D task with O(100) electrodes), merely recording more neurons without other improvements in parallel may not lead to substantial further performance gains on simple 2D reaching tasks. This potential trend of saturation could be due to the following reasons:
First, we dropped the channels in reverse order of importance defined by mutual information between the spike distributions and reach direction. Thus, the decoder performance is less likely to improve when the added channels are less informative of hand kinematics. Second, it’s possible that the behavioral task here is relatively simple and does not require more neurons to reach high performance. We note that for a more complex task, increasing the number of channels could further improve rather than saturate the performance. Third, the current decoder algorithm might be unable to improve the performance as the channel number increases beyond a threshold; this motivates the need to develop more advanced decoders [22]. Decoding many thousands of neurons recorded via calcium imaging in NHPs may enable us to better develop and evaluate the performance of new decoder algorithms as a function of the number of neurons.

B. Robustness to Gaussian Noise

We evaluated the decoder’s performance on the original spike data, and on the surrogate calcium imaging data generated with different levels of Gaussian noise. We found that the decoder had better performance, as indicated by a lower MSE and higher Pearson’s r, for 1000 Hz imaging data compared to spike train data, when the noise level was low (Fig. 3A). (Importantly, we do not purport to ascribe specific physical source or interpretation to a given level of noise; it is the trends that we focus on.) This may well be due to the smoothing of data by the exponential decay kernel convolution (equation 5 or 6), which could in principle improve offline decoder performance. Moreover, the decoder showed some robustness to time-invariant Gaussian noise added to the surrogate data, as shown in Fig. 3A. The performance was still comparable to that of spike data even when \( \sigma \) was 10 times \((\Delta F/F)_{\text{unitary}}\), but dropped as the noise level was further increased (Fig. 3A). This also served as a basic check that simulations were working properly, as increasing noise beyond some level must erode decoder performance. Saturation of the fluorescence signal, but not signal rise time, also consistently worsened decoder performance (Fig. 3A).

C. A Realistic, but Still Fairly High, Sampling Rate (200 Hz)

Current in vivo 2p imaging can achieve frame rates up to approximately 200 Hz [19], though the average frame rate of most studies is 30-60 Hz or lower for the purpose of, for example, higher SNR and larger fields of view [18], [23]. We first evaluated the decoder’s performance at the sampling rate of 200 Hz. As expected, down-sampling reduced decoder performance, with larger MSE and smaller r (Fig. 3B). However, when sampling at 200 Hz, we found that the decoder could still decode hand movement directions and hand positions reasonably well (example decoded trajectories in Fig. 2B). Incorporating the rise time and non-linearity of calcium indicators into the down-sampled datasets further worsened decoder performance (Fig. 3B). Down-sampling to 100 Hz or lower rates (for instance, 30 Hz in Fig. 3B) led to a substantial decline in BMI decoder performance, as indicated by MSE and Pearson’s r measures, even for a low-noise surrogate dataset \((\sigma'(\Delta F/F)_{\text{unitary}}=1)\). At a 100 Hz sampling rate, we observed an MSE of 10528 and Pearson’s r of 0.53 for the synthetic imaging data with a 45-ms rise time and saturation, while at a 30 Hz sampling rate they were 13798 and 0.40, respectively.

IV. DISCUSSION

Here we investigated, by decoding surrogate calcium imaging data, how optical imaging signals could provide information about movement kinematics. We found that a Kalman filter based decoder (FIT-KF) could accommodate some degree of Gaussian noise that was added to the initial surrogate optical data to more accurately simulate 2p calcium imaging signals. Part of this benefit could be due to the summing of neural activity over a bin by the decoder that helped overcome some of the normally-distributed noise with 0 mean. We also demonstrated that our decoder performance was quite sensitive to sampling rate, with MSE doubling when going from a sampling rate of 1000 Hz to 200 Hz. The combination of rise time and saturation of fluorescence signals further decreased the decoder performance at this sampling rate. However, movement directions and trajectories could still be decoded to a large extent. Sampling rates at 100 Hz or lower (for instance, the commonly-applied 30 Hz) dramatically impaired decoder performance. This suggests the need for improvements in imaging techniques and decoding algorithms.

Interestingly, we found that at a relatively low noise level \((\sigma'(\Delta F/F)_{\text{unitary}}=1)\), the decoder performance was seemingly

<table>
<thead>
<tr>
<th>TABLE I. Factors influencing Imaging Signal Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sources of noise and non-linearity</td>
</tr>
<tr>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Fluctuation of baseline calcium concentration</td>
</tr>
<tr>
<td>Various amounts of calcium influx for each spike</td>
</tr>
<tr>
<td>Buffering of intracellular calcium that doesn’t contribute to fluorescence signals</td>
</tr>
<tr>
<td>Non-linearity of calcium-indicator binding (in particular, cooperative binding)</td>
</tr>
<tr>
<td>Saturation of calcium indicators at high calcium concentration levels</td>
</tr>
<tr>
<td>Variability in calcium dynamics across neurons</td>
</tr>
<tr>
<td>Number of released photons by the calcium indicators roughly following Poisson process</td>
</tr>
<tr>
<td>Optical collection and measurement noise in microscope imaging systems</td>
</tr>
<tr>
<td>Motion of imaging plane</td>
</tr>
</tbody>
</table>

a. Incorporated into the Gaussian noise added to simulated fluorescence signals, but not explicitly estimated with a physiological model. b. Here we used the saturation model for calcium dyes such as OGB-1 (see Methods).
better for the non-down-sampled calcium data than for original spike train data. A putative explanation is that the fluorescence signal emitted from calcium indicators has as an inherent part of it a simple (biochemical) smoothing low-pass filter and thereby can help offline decoder performance. This natural smoothing could be a benefit as noted, but could also limit performance by increasing the closed-loop latency [24].

In summary, we report here, to our knowledge, the first feasibility and design study for o-BMIs in NHPs. We expect that decoder performance will improve with increasing numbers of recorded neurons, which is a major advantage of imaging techniques, although performance tends to saturate if we merely increase neuron numbers without other improvements. Our results also provide insight into how to choose imaging parameters (e.g. optical frame rate, optical scan path, and laser power that influence temporal resolution and SNR) to optimize decoder performance. Importantly, the imaging parameters we found leading to reasonable decoding performance may require additional advances in recording technology, because of the trade-off between frame rate and SNR. Besides improving microscope systems, developing faster, more linear calcium or voltage indicators will further enhance the quality of imaging signals [25], [26], in addition to optimization of o-BMI algorithms for optically recorded data. We note that when converting spikes to the surrogate fluorescence signal, our model did not address the difference in recorded neural populations by electrophysiology and imaging methods. Specifically, electrodes often detect multiunit signals and can reach deeper cortical layers whereas imaging measures single neuron activity and is currently limited to superficial layers. We also note that o-BMI approach has not yet been tested in online, closed-loop BMI experiments which is critical – above and beyond offline analyses – as only in closed-loop can controllability and performance be better evaluated [7], [24]. Furthermore, there are other sources of noise that can corrupt real imaging data but were not incorporated into our model, including motion of the imaging plane, drift of baseline fluorescence, cell-to-cell variability, and other potential factors (Table I lists these many considerations) [27]. We note that much of this noise can be reduced by data pre-processing algorithms such as image registration and across-trial averaging [28], [29]. We believe that it will be possible to overcome many of the present-day challenges, and embrace the many o-BMI opportunities, through principled end-to-end system design.

ACKNOWLEDGMENTS


REFERENCES


