Proceedings of the 2nd International Conference on Modern Medicine and Global Health
DOI: 10.54254/2753-8818/45/20240759

Brief analysis of procedure in optogenetics

Zhaoxi Chen

Department of Brain and Cognitive Science, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, United States

zhaoxic2@illinois.edu

Abstract. This article aims to provide a brief analysis of the methods utilized in optogenetics data analysis procedure. The two main categories of procedure are online and offline analysis methods, which determines the causality of the system. Both approaches have their advantages and tradeoffs, and this article will introduce some examples of each procedure in an experimental setting, what are some specific processing methods they are using, what metrics are explained, how the two approaches' performance compare, and how they produce and represent analytical data. One practical example on each approach, performed on a dataset from a previous experiment imaging neuronal activity in rats, is provided for better clarity and comparison measures. The offline example fits gaussian mixture model on high-pass filtered data, and the separate gaussian models are used to predict the status of the neuron given its relative fluorescence intensity. The online example uses an existing algorithm called OASIS, which uses an autoregressive model to reconstruct calcium trace and thus infers the spike. The main difference between the two is demonstrated to be their robustness and accuracy: online approach is more robust and can be utilized while recording the data, giving interpretable results with low latency, yet its accuracy does not depend on obtained sample number; offline approach is more time-consuming while fitting and training data to an optimal model. However, offline approaches' accuracy will increase with large sample size. Both approaches provide deep insight into the acquired datasets, and while analyzing data they should be used strategically to fit the specific needs of the task.

Keywords: Optogenetics, Calcium Imaging, OASIS

1. Introduction

Optogenetics has been one of the most discussed and explored topics of bioimaging in the last few decades [1]. Its unique advantages, such as the feasibility of in vivo recording of fluorescence signal from numerous neurons simultaneously, provide an indirect measurement of neuronal activity with high spatial resolution [2,3]. By measuring fluorescence induced by calcium influx through genetically modified calcium receptors, an indirect measure of neuronal activity is provided [4]. However, an indirect measure will need to be processed in order to infer the original neuronal activity, which is the aim of this experiment project. In this project, we first pre-process the data, then categorize filtered data into two states offline, and calculated the connectivity between each neuron in each dataset. Ultimately, we use an online deconvolution algorithm called OASIS and compared its results to ours. In the discussion section the two methods are then compared in terms of efficiency and accuracy.

DOI: 10.54254/2753-8818/45/20240759

2. Methods

The original datasets contain three variables: frame rate, which is essentially sampling rate; location of the neurons; neurons and their relative fluorescence intensity across time. By taking each neuron's recorded activity and Fourier transforming the activity in the time domain into the frequency domain, it is clear that the raw data has many low-frequency artifacts, which may be due to motion or photo bleaching. By filtering out these artifacts, we can increase the signal-to-noise ratio. A high-pass filter is designed to filter out the low-frequency noise by a cut-off frequency, which is identified from the first large peak with relatively high frequency in the frequency domain. After we successfully filter out the noise in the frequency domain signal by taking the dot product of the Fourier-transformed signal and the filter, we inverse-Fourier-transform the filter into the time domain. Unwanted low frequency noise in the raw data is filtered out from the time domain by convolving the filter with the raw data in the time domain.

After filtering the raw data, a two-factor gaussian mixture model is fitted (using a generic function in Matlab) on each neuron across all corresponding samples. This process creates a probability model that can distinguish between the resting state and excited state of a neuron and can be used to classify and predict a neuron's state given filtered data. One method is to calculate the p-value of the neuron in each state (for each gaussian distribution) and compare the p-values, taking the larger p-value. Then the correlational coefficient of each neuron pair, as their connectivity metrics, is computed using the corrcoeff() function from Matlab and drew a color-scaled connectivity graph for each dataset. This connectivity test could tell us how the fluorescence pattern of each neuron pair is related, such that there exists a positive or negative relation, or no relation at all.

The final step is to use OASIS and its deconvolveCa() function for analysis of the raw data. This involves two steps in the algorithm, which are reconstructing calcium trace and inferring spikes (deconvolution). We assume that filtration of the raw data is not needed as the autoregressive process can directly filter out lower frequency artifacts such as movement and higher frequency artifacts such as microwave radiation. Calcium trace reconstruction involves the OASIS algorithm, which is for isotonic regression. Its mechanism can be described as:

First, compare the current data point with the next data point. If the relative intensity of the next data point is larger than that of the current data point multiplied by some constant b, move to the next data point. If not, it is seen as a violation of the predetermined rules. Then the algorithm would operate backwards until it encounters the most recent spike, essentially calculating the fluorescence decay in an exponential trace, as the same constant was applied multiple times [2].

After reconstructing the calcium trace, the calcium trace is then used to infer the spikes. As Stringer & Pachitariu suggest, spikes can be considered as the filtered signal, and fluorescence decay in calcium traces as noise to be filtered. The process of inferring spikes is then converted to a process of convolving a filter on the calcium trace in the time domain, or deconvolving an inverse filter [5]. The spike time can be easily inferred from the sudden increase in fluorescence, and the spike intensity can be inferred from the difference between the intensities of the calcium trace spike and the decay after it. From the spike timing, we can calculate the spike rate of each neuron, which is our ultimate target.

3. Results

The average Fourier magnitude spectrum of dataset 1 shows a peak at around 0.198, whereas the spectrum of dataset 2 shows a peak at around 0.187. To preserve the first peak for both datasets, a cut-off frequency of 0.18 is chosen to maximize filter effect. The result and comparison with raw signal are shown in Figure 1. The filtered signal is then inputted into the fitgmdist() function to create gaussian mixture models for each of the neurons. The probability density function of neuron 1 in dataset 1 is shown in Figure 2. From the probability distribution we observe a large overlap between the two gaussian distributions for two neuronal states, yet it is clear when relative intensity is around zero the probability of predicting the state of this neuron to be excited is far greater than predicting that it is in resting state. If we plug in the specific intensity we will be able to calculate the conditional probability for both states and obtain a decision.

DOI: 10.54254/2753-8818/45/20240759

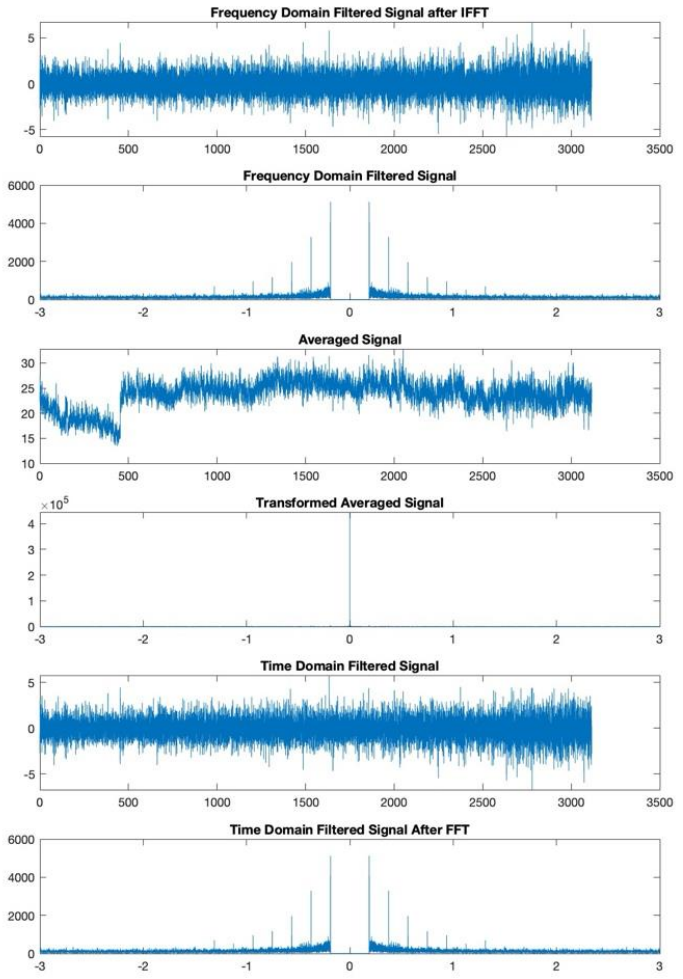


Figure 1. Comparison of raw signal (graph 3, 4) and signal filtered in frequency domain (graph 1, 2) and time domain (graph 5, 6). Graph 1, 3, 5 is the specific signal plotted in time domain; Graph 2, 4, 6 is the specific signal plotted in frequency domain.

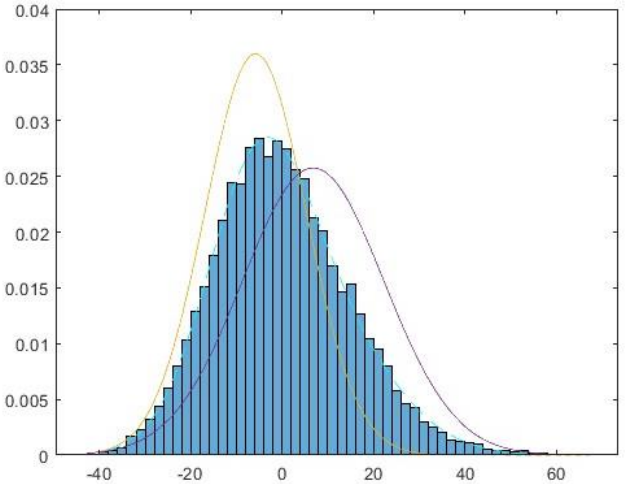
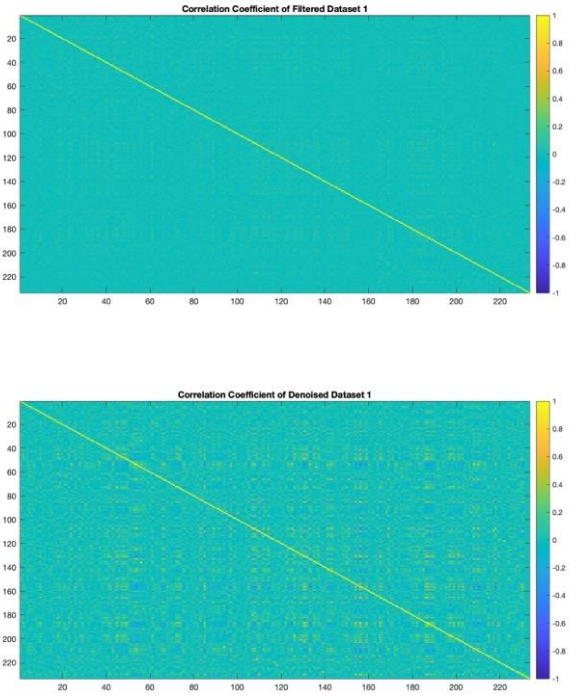
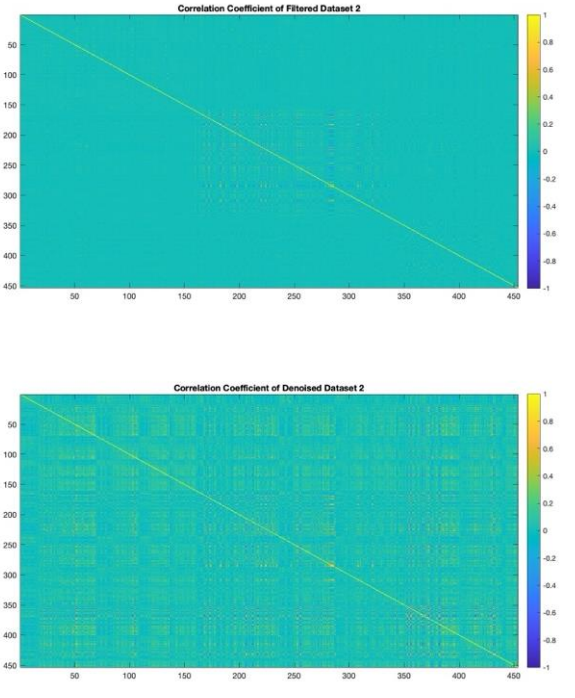


Figure 2. Probability density distribution of the state of a neuron. The x-axis is the relative intensity in the data, and the y-axis is the probability. The yellow line and purple line show separate gaussian distribution for the states of the neuron. Yellow is resting state, and purple is excited state.

DOI: 10.54254/2753-8818/45/20240759



(a)



(b)

Figure 3. Color-scaled correlation coefficient matrix from both datasets and both filtered data and calcium trace. Figure 3a. on the left side shows comparison of correlation coefficient matrix between filtered data (upper graph) and calcium trace (lower graph) from dataset 1. Figure 3b. on the right side shows comparison of correlation coefficient matrix between filtered data (upper graph) and calcium trace (lower graph) from dataset 2.

DOI: 10.54254/2753-8818/45/20240759

The comparison between correlation coefficient matrices is shown in Figure 3. We can observe that there are stronger correlations between neurons in calcium trace correlation coefficient matrices. This may partly be due to the OASIS algorithm filtering out the noises by the autoregressive approach and thus has higher signal to noise ratio.

Using deconvolveCa() function from the OASIS package, we successfully reconstructed the calcium trace and inferred the spike timing from the raw data. The result of deconvolveCa() function on neuron 1 from dataset 1 is shown in Figure 4. After inferring the spikes, we had access to spike count, which is essential for calculating spike rate, as spike rate is spike count divided by the length of sample time. The spike rate from the first 30 neurons in dataset 1 is shown in Table 1. The spike rate ranges from about 0.05Hz to around 3Hz, which is reasonable for neurons, yet the lower range may be of concern.

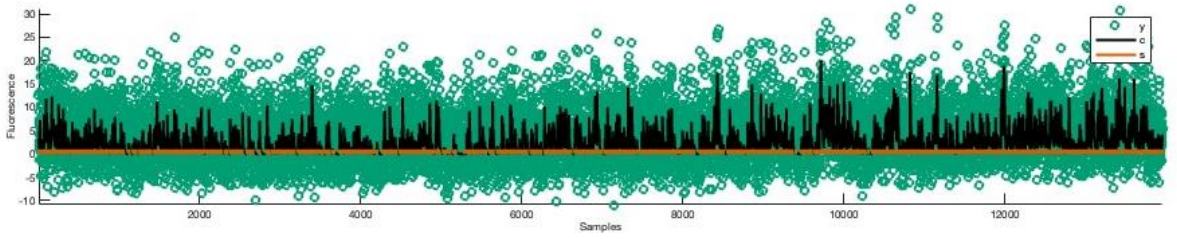


Figure 4. Result of deconvolveCa() function on neuron 1 from dataset 1. X-axis is sample frame. Y is raw data points, with y-axis representing relative intensity. C is calcium trace. S is spikes inferred.

Table 1. Spike rate of the first 30 neurons in dataset 1. Row 1 contains neurons 1 to 10, row 2 contains neurons 11 to 20, row 3 contains neurons 21 to 30.

0.8921	0.8602	0.5338	0.6118	0.2680	0.7777	0.2211	0.7549	0.3519	2.1270
0.4322	0.3446	0.2357	0.5438	0.4486	0.8721	0.3770	0.9660	1.1602	0.6870
1.1387	0.2557	2.6125	0.9619	0.6191	2.0705	0.6975	0.6915	0.7937	0.2603

4. Discussion

OASIS provides a comprehensive and time efficient procedure for reconstructing action potential spikes. It is developed as a fast online data processing method for optogenetics using calcium trace imaging. An offline approach, such as the procedure of gaussian mixture model in our method, is then time consuming and purely causal comparing to OASIS. In this perspective OASIS might provide a more robust approach for data processing with only future inputs, requiring no previous data, yet it might fail to provide a higher accuracy compared to the result from offline approaches with large data. Further research may target on effect of larger dataset to the accuracy of offline methods, and how the efficiency compares to online methods.

5. Conclusion

Optogenetics with calcium imaging technique requires appropriate data processing procedure to explain the data, and the two categories of approaches: whether online or offline, are purpose-specific yet powerful methods. They have their unique tradeoffs in term of temporal resolution and accuracy, and depending on the situation one might outperform the other in some means. OASIS is a robust deconvolution algorithm with low latency and is suitable for online use, yet if time budget is not a concern an offline prediction algorithm with higher accuracy, such as gaussian mixture model, should be of higher priority. Some further research that could be done on basis of this finding may be to combine the advantages from both approaches and optimize for an algorithm with low latency and high accuracy. An artificial-neural-network-based algorithm with higher prediction accuracy than existing offline algorithms concluded from large sample size may also be of interest for further investigation.

References

[1] Deisseroth, K. (2010). Optogenetics. *Nature Methods, 8*(1), 26–29. doi:10.1038/nmeth.f.324

DOI: 10.54254/2753-8818/45/20240759

[2] Friedrich, J., Zhou, P., & Paninski, L. (2017). Fast online deconvolution of Calcium Imaging Data. *PLOS Computational Biology*, *13*(3). doi:10.1371/journal.pcbi.1005423

[3] Pnevmatikakis, E. A. (2019). Analysis pipelines for calcium imaging data. *Current Opinion in Neurobiology, 55*, 15–21. doi:10.1016/j.conb.2018.11.004

[4] Grienberger, C., & Konnerth, A. (2012). Imaging calcium in neurons. *Neuron, 73*(5), 862–885. doi:10.1016/j.neuron.2012.02.011

[5] Stringer, C., & Pachitariu, M. (2019). Computational processing of neural recordings from Calcium Imaging Data. *Current Opinion in Neurobiology*, *55*, 22–31. doi:10.1016/j.conb.2018.11.005