Tracking Neural Activity: Automated Image Analysis

BethAnna Jones, Amanda Stanley

July 19, 2018

Abstract

This work was completed as part of the 2018 SURIEM program supported by Michigan State University, the National Security Agency, and the National Science Foundation. Modern optical methods such as high-definition two-photon imaging of calcium fluorescence allow us to view the activity of tens of thousands of neurons in the brain. However, due to the large number of images to sort through, manual methods of location, identification, and analysis of these individual neurons is tedious and time-consuming. Researchers need an automated system to complete in hours what may take an expert weeks or even years.

Some automated techniques exist, but as shown by the 2017 SURIEM project, they find only roughly half the cells, and disagree on half of what they find. We aim to build on this work and improve automatic analysis techniques of neural images. Current leading methods analyze neural images using singular value decomposition and constrained non-negative matrix factorizations. We will compare the cells identified with these methods and improve them by i) analyzing fluorescence patterns of the cells; ii) investigating correlation patterns across cells; and iii) aligning cells between sessions in order to compare cell activity across days as animals learn.

This work was completed as part of the 2018 SURIEM program supported by Michigan State University, the National Security Agency, and the National Science Foundation.

1 Introduction

Modern optical imaging techniques are poised to revolutionize neuroscience by enabling the processing and analysis of large amounts of data. If we wish to understand how brain activity correlates with specific actions, we must first understand how individual neurons behave and interact with one another. High quality videos of the brain allow us to view brain activity on the neural level.

We have been able to investigate strategies and methods of analyzing videos of rodent brains by working with about a dozen labs around the world, including Stanford and Lethbridge. Leading research in optical images of mice brains allows neuroscientists to detect and track firing neurons and their correlation to mouse behavior.

However, it is very hard to identify individual cells in images. Furthermore, one of the major goals of neuroscience is to observe how the activity dynamics of individual cells and circuits change during learning. That is, study the changes in circuit dynamics as animals remember specific events or become more familiar with specific locations, objects or tasks. Very few labs have managed to record activity from a large number of the same neurons over many days.

Our two goals for this year's summer project were:

1. Try to build better models of mathematical and statistical models of fluorescence related to cell activity
2. Try to develop a semi-automated method for aligning cells in two different imaging sessions.

2 Models of Cells

2.1 Two-photon Imaging and Calcium Fluorescence

The transgenic mice studied in these experiments produce GCaMP, a genetically-encoded calcium indicator. This indicator allows for neural videos to be created through a process known as two-photon imaging.

Using a focused laser of two photons, this form of fluorescence imaging allows for creation of images up to about 1mm depth of the brain. The calcium indicators fluoresce at a variety of levels of intensity when being released from neurons firing and is picked up by a high-sensitivity detector. Each focal point that the laser detects becomes a pixel in the image where each pixel is roughly 1μm. Scanning a typical 300×300 pixel area using a two-photon imaging laser over a span of time creates videos of planes of illumination of the brain. Using the fact that neurons are approximately 10μm in diameter, these videos are later broken down through algorithms to locate neurons.

During the process of two-photon imaging, two photons of near-infrared excitation light at long wavelengths are directed at the selected tissue area. The photons cause the fluorescent dyes at the focal point to be excited and emit a single, fluorescent photon. A photomultiplier tube collects this photon, and allows for the focal point area to be captured as a pixel. Not only does employing the two-photon method allow for thinner imaging planes to be captured and in greater detail than from other leading methods (such as single-photon excitation microscopes), but two-photon imaging also causes less damage to the living cells.

2.2 Matrix Factorization

We represent a data set (a video of the brain) as a 3-dimensional matrix $\mathbf{B} \in \mathbb{R}^{N \times D_y \times D_x}$ where

$$\begin{aligned} N &= \text{number of frames} \\ D_y &= \text{number of pixels along the } y\text{-axis of the video} \\ D_x &= \text{number of pixels along the } x\text{-axis of the video.} \end{aligned}$$

Note that $\mathbf{B}$ is very large. In order to make these data sets more manageable, we convert to a rectangular array with dimensions $N \times D$ where $D = D_y \times D_x$. For example, for a short video $\mathbf{B} \in \mathbb{R}^{4\times5\times4}$, we can reshape $\mathbf{B}$ into the rectangular matrix $\mathbf{A} \in \mathbb{R}^{4\times20}$. We continue with the following alternate definition of a neuron.

Definition 2.1. Let a **source** be defined by a signal and a footprint. A source's **signal** is a measurement of its relative activity over time, and a source's **footprint** is its spacial measurement (location, size, and shape).

Note that we will use the terms neuron and source interchangeably. We can now use different methods of dimension reduction to work with these high-dimensional arrays and find our sparse set of k sources. We aim to find these true sources in our data and eliminate background noise caused by **neuropil**, the dense network of glia, neurons, dendrites, and axons in the brain. Let

- $\mathbf{S} \in \mathbb{R}^{N\times k}$ be the signal data for the k sources
- $\mathbf{F} \in \mathbb{R}^{k\times D}$ be the footprint data for the k sources
- ξ be the "noise" error created by the neuropil.

Then for k sources we use **matrix factorization** to give us

$$\begin{aligned} \mathbf{A} &= \mathbf{SF} + \xi \\ &= \sum_{m}^{k} \mathbf{S}_m \mathbf{F}_m + \xi. \end{aligned}$$

Two common algorithms used to identify cells within these videos are Suite2P and Calcium Imaging Analysis, or CaImAn. These algorithms use different methods of matrix decomposition to pick out neurons: singular value decomposition (SVD) and nonnegative matrix factorization (NNMF), respectively. Both methods first require reshaping the data from a three-dimensional matrix into a two-dimensional matrix.

For our project, NMF is used to factor the reshaped video matrix into a tall, thin matrix $\mathbf{W}$ and a short, long matrix $\mathbf{H}$, where $\mathbf{W}$ is an estimate of $\mathbf{S}$ and $\mathbf{H}$ is an estimate of $\mathbf{F}$.

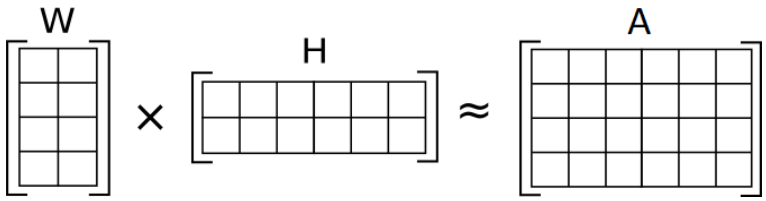


Figure 1: Visual representation of nonnegative matrix factorization

Note $\mathbf{H}$ can then be reshaped back to a three-dimensional matrix with dimension $k \times D_y \times D_x$. This should be an estimate of the original video (with the signals removed) where it should be easier to identify neurons.

The second algorithm, Suite2P, uses singular value decomposition, another form of matrix decomposition.

Theorem 2.1 (**Singular Value Decomposition**). *Let* $\mathbf{A} \in \mathbb{R}^{m \times n}$. *Then there exists a factorization* $\mathbf{A} = \mathbf{U\Sigma V^T}$ *such that*

- $\mathbf{U}$ *is an* $m \times m$ *orthogonal matrix,*
- $\mathbf{\Sigma}$ *is an* $m \times n$ *diagonal matrix of nonnegative numbers,*
- $\mathbf{V}$ *is an* $n \times n$ *orthogonal matrix.*

The diagonal entries of $\mathbf{\Sigma}$ *are known as the* ***singular values*** $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ *of* $\mathbf{A}$ *where* $r = min(m, n)$.

For example, if $\mathbf{A} = [a_{ij}]$, then there exist orthogonal matrices $\mathbf{U} = [u_{ij}]$ and $\mathbf{V} = [v_{ij}]$ and diagonal matrix $\mathbf{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$, such that

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}.$$

Similar to NMF, the three-dimensional video matrix $\mathbf{B}$ must first be reshaped into a two-dimensional matrix $\mathbf{A}$ before it can be factored.

Consider that the factors that contribute most to the activity and fluctuations in our data set will be from the neurons we are trying to identify. Let $\mathbf{u_i}$ and $\mathbf{v_i}$ be the column vectors of $\mathbf{U}$ and $\mathbf{V}$ respectively. Then

$$\mathbf{A} = \sum_{1}^{r} \sigma_i \mathbf{u_i} \mathbf{v_i}^T \tag{1}$$

$$= \sigma_1 \mathbf{u_1} \mathbf{v_1}^T + \sigma_2 \mathbf{u_2} \mathbf{v_2}^T + \ldots + \sigma_r \mathbf{u_r} \mathbf{v_r}^T \tag{2}$$

Note that the ith singular value σ_i acts as a scalar weight for the ith contributing factor $\mathbf{u_i} \mathbf{v_i}^T$. Hence, by identifying the largest singular values in Σ, we identify the factors that contribute most to the reconstruction of $\mathbf{A}$, and by adding only those up, we eliminate most of the noise created by neuropil and find an

approximation of A,

$$\hat{\mathbf{A}} = \sum_{1}^{r} \sigma_i \mathbf{u_i} \mathbf{v_i}^T \tag{3}$$

$$= \sigma_1 \mathbf{u_1} \mathbf{v_1}^T + \sigma_2 \mathbf{u_2} \mathbf{v_2}^T + \ldots + \sigma_k \mathbf{u_k} \mathbf{v_k}^T. \tag{4}$$

The following definition allows us to more easily find the largest singular values.

Definition 2.2. We define the **variances** of the singular value decomposition of a matrix **A** to be the squares of its singular values.

Variances are often used in statistical analysis, but also allow us to more clearly find distinctions between the singular values. From Figure 2, we're able to estimate that the singular values begin to become insignificant when $i = 5$.

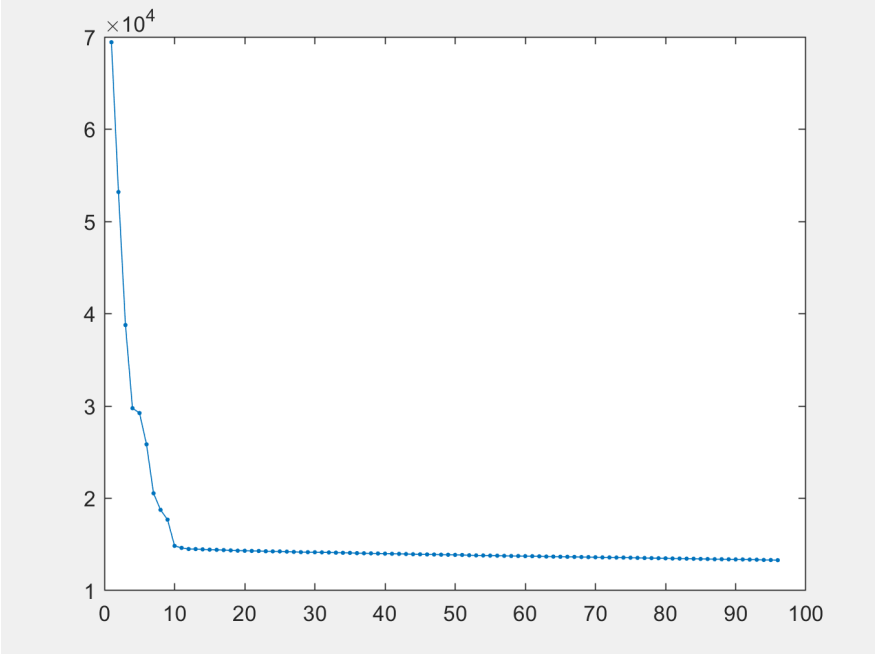


Figure 2: The variances of our synthetic data.

Let $k = 5 \leq r$. From Equation (4),

$$\begin{aligned}\mathbf{A} &= \sum_{i}^{r} \sigma_i \mathbf{u_1} \mathbf{v_i}^T \\ &= \sigma_1 \mathbf{u_1} \mathbf{v_1}^T + \sigma_2 \mathbf{u_2} \mathbf{v_2}^T + \ldots + \sigma_r \mathbf{u_r} \mathbf{v_r}^T \\ &= \sigma_1 \mathbf{u_1} \mathbf{v_1}^T + \sigma_2 \mathbf{u_2} \mathbf{v_2}^T + \ldots + \sigma_k \mathbf{u_k} \mathbf{v_k}^T + \ldots + \sigma_r \mathbf{u_r} \mathbf{v_r}^T.\end{aligned}$$

We can thus find an approximation of our data,

$$\begin{aligned}\tilde{\mathbf{A}} &= \sum_{1}^{k} \sigma_i \mathbf{u_i} \mathbf{v_i}^T \\ &= \sum_{1}^{5} \sigma_i \mathbf{u_i} \mathbf{v_i}^T \\ &= \sigma_1 \mathbf{u_1} \mathbf{v_1}^T + \sigma_2 \mathbf{u_2} \mathbf{v_2}^T + \ldots + \sigma_k \mathbf{u_k} \mathbf{v_k}^T.\end{aligned}$$

Our approximation $\tilde{\mathbf{A}}$ will have less background noise and allow us a clearer view of the neurons:

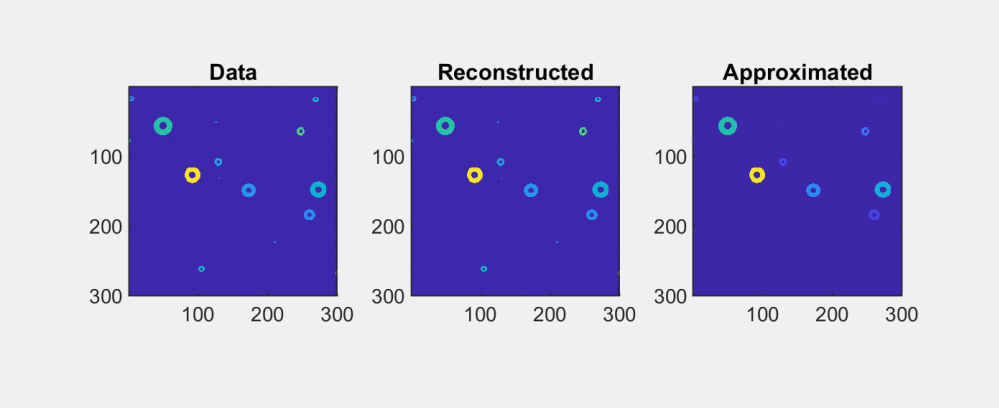


Figure 3: Cell Finding Using Singular Value Decomposition

3 Alignment and Matching Sessions

A major goal of neuroscience is to study how brain activity changes as an animal learns or becomes familiar with a task. For this we need to record the same cells on different days. However it is difficult to set the imaging plane exactly in the same place. Therefore we try as best we can and try to align images afterward.

Most cells look somewhat alike. Our approach combines geometry and activity signatures. The biggest geometric features are the blood vessels: most transverse venules and capillaries look very similar in cross-section, whereas often the few capillaries running partly within the imaging plane are highly distinctive.

Independent combinations of features in the same spatial relationships—to a tolerance of a few microns—are highly unlikely to reoccur in a random slice in roughly the same area. We expect some cells' activity profiles to change from day to day. However we expect the majority of cells to maintain their characteristics. This will enable us to validate matches when activity profiles have a much higher correspondence than expected by chance.

We prepared a number of sessions for alignment and aligned up to 10% of cells. Using three geometric features that are the same in two planes, the alignment process used a graphical user interface to transform the regions of interests from the two planes, matching cells identified in both planes.

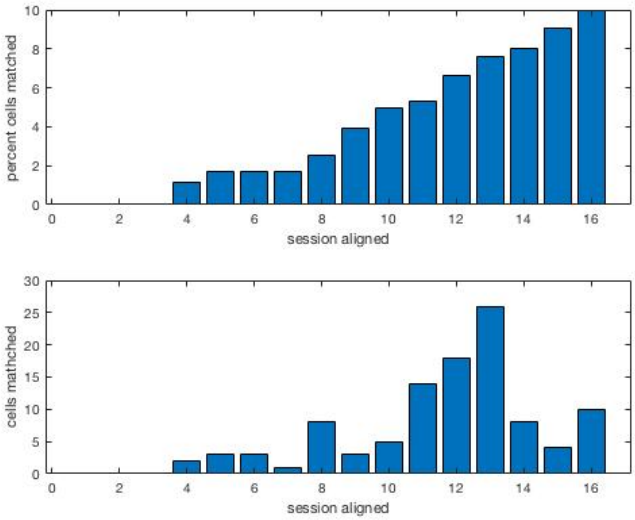


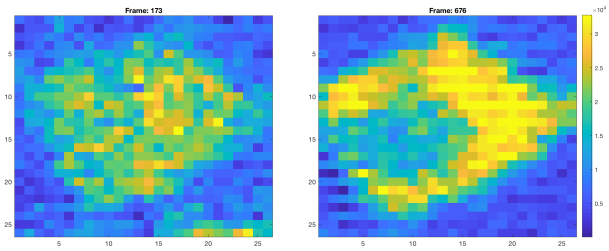
Figure 4: Cell alignment data

Observing Figure 4, we see that three of the sixteen aligned sessions did not match any cells. This was most likely due to the lack of visual geometric features. However, we did not have time to compare the activities of the same cells on different days.

4 Cell Substructure

Most current models assume that cell activity gives rise to fairly consistent dynamics and rises and falls in calcium fluorescence have very similar time courses over the whole geometric extent of the cell. This fits the factor analysis approaches to identifying individual cells in the image stacks very well. Departures from this ideal are assumed to be random and to reflect only optical noise. However, so far these models have not worked very well; at best the models reduce the amount of time spent manually annotating cells. But they do not seem likely to scale to the massive data sets expected over the next few years. However, if patterns of calcium fluorescence are systematically different in major ways, these differences could be explicitly modelled to improve the estimates of cell activity.

Peak Brightness Throughout the various stages of the firing process of a neuron, the calcium indicators fluoresce in particular patterns. We aimed to study these patterns and cell substructure in order to improve the accuracy of future neuron-finding algorithms. Typically, as a neuron receives a signal, it fluoresces near the center of the cell and then the wave of fluorescence heads towards its axon and then towards its dendrites as the signal is propagated.



For each region of interest in the neural video, we selected the individual frames in which the cell was at "peak brightness." For each region of interest, we selected a rectangle of pixels (that outlined the region of interest) and found the frame in which the magnitude of the fluorescence was at its greatest. Since individual neurons have varied maximum magnitudes of florescence, we considered each region of interest individually.

Once this frame was selected, we we able to more closely examine some of the patterns of florescence.