AUTOMATED IDENTIFICATION OF NEURONAL ACTIVITY FROM CALCIUM IMAGING BY SPARSE DICTIONARY LEARNING

Ferran Diego[1] *Susanne Reichinnek*[2] *Martin Both*[2] *Fred A. Hamprecht*[1]

[1]HCI, IWR, University of Heidelberg, 69115 Heidelberg, Germany
[2]Institute for Physiology and Pathophysiology, University of Heidelberg, 69120 Heidelberg, Germany

ABSTRACT

We present ADINA, an automated pipeline for analyzing and identifying neuronal activity from calcium imaging data to investigate neuronal activity patterns. This entails the detection and classification of cell centroids and of calcium transients (events) that reappeared during different activity periods as memory consolidation. Specifically, the pipeline implements a sparse dictionary learning to infer the most relevant Ca^{2+} patterns, an image segmentation procedure using a wavelet–transform and watershed to identify single cells, and an estimation of the transient signals by means of sparse coding exploiting spatial and temporal sparsity. We validate our automated approach on artificial and two different calcium imaging sequences from mice hippocampal slice cultures acquired with fluorescence and confocal microscopes. Our approach achieves ca. 94% sensitivity on average for correctly detecting events, thus improving significantly the estimation of cell signals relative to published procedures.

Index Terms— Calcium Imaging, Cell Sorting, Dictionary Learning, Sparse Coding, Segmentation

1. INTRODUCTION

Calcium imaging is an increasingly popular technique for monitoring simultaneously the neuronal activity of hundreds of cells at single cell resolution. This makes it an essential tool for studying complex patterns of distributed activity in neuronal networks of local circuits of the brain like decision making or conciousness (Fig. 1). However, as [1] noted, the extraction of this information is hampered by several complications: 1) low signal-to-noise ratio (SNR); 2) uneven background; 3) misleading fluorescence signals due to strong light scattering; and 4) overlapping neurons.

In addition to the challenges described above, computational techniques for rapid and reliable extraction of neuronal activity patterns from complex 4-dimensional (space and time) data sets still need major improvement [1]. Most approaches involve the delineation of regions of interest (cells, or parts of cells) by eye [2] or in a semiautomated fashion [3]. However, this procedure is subjective and does not scale to large–scale imaging data. Several algorithms have been proposed in the past to assign local calcium transients to distinct cells based on machine learning [4], matrix decomposition [5, 6] or image processing techniques [1]. All these works focus mainly on detecting the location of cell centroids on which transient signals are extracted from the raw data over time, and are dependent on a specific imaging technique. For instance, Maruyama *et al*. [6] and Mukamel *et al*. [5] decompose calcium imaging data into constituting signal sources, that is, spatial and temporal components. The former uses non–negative matrix factorization (NMF) which cannot discriminate cells that are highly correlated in the temporal domain. Hence, the latter combines principal component analysis (PCA) and independent component analysis (ICA) with a subsequent image segmentation to extract single cells; however, its performance suffers when the point spread function is very anisotropic, as in fluorescence microscopy.

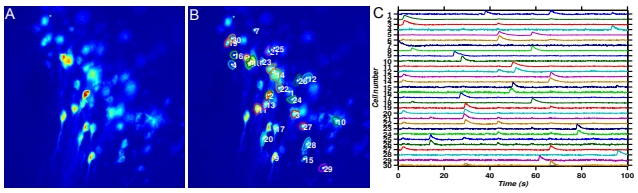

Fig. 1. Identification of neuronal activity. a) Maximum intensity projection across a sequence of 1000 frames using false color plots. b) Centroids and contours of 30 detected cells and c) the respective Ca^{2+} transient signals.

Therefore, in this paper, we present an automated pipeline for monitoring the activity of multiple cells from calcium imaging data recorded by different microscopes, *e.g*. confocal or epi–fluorescence (Fig. 2). We focus on the detection and classification of cell centroids and of cell shape ("*cell sorting*") (Fig. 1b), and the inference and analysis of Ca^{2+} transients signals (events) ("*transient analysis*") (Fig. 1c). The extraction of cellular signals from calcium imaging data is very challenging since calcium events do not fit individual cell identities and also single image pixels may represent more than one distinct cell or noise. Although the cell identities have non–specific appearance and shape, we can exploit that these events are sparsely distributed in both space and time. Hence, our *cell sorting* relies on a flexible matrix factor-

Ferran Diego would like to thank the Excellence Cluster "Cellular Networks" for financial support.

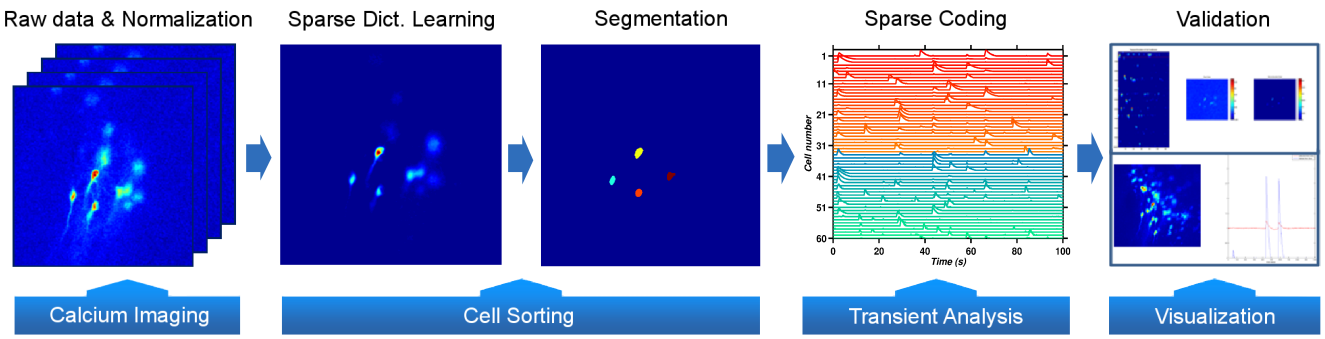

Fig. 2. Workflow for identifying neuronal activity. Sparse dictionary learning produces a set of basis functions to describe the spatial patterns of single or grouped cells. These basis functions are segmented into unique single cells for which transient signals are inferred by sparse coding. The final cell sorting and transient signals are visualized together with raw data for validating the results.

ization based on sparse dictionary learning [7] that exploits this sparsity, allows more flexibility to adapt the representation to the data and does not impose orthogonality of the basis functions. This is combined with a versatile image segmentation that improves the detection of cell shape.

Our *transient analysis* infers the transient events using sparse coding [7] to exploit both the spatial information of the extracted cells and cell sparsity in space and time. This is in constrast to [1, 4, 5, 6] who use only the raw data extracted from the cell centroids. Inferring transient signal has the advantage that these events are inferred accurately and makes it easier for detecting and interpreting visually when cells spike or not than inspecting directly raw data; furthermore it is computationally efficient given the segmented cells.

2. AUTOMATED CELL DETECTION

2.1. Workflow Overview

The identification of cellular activity relies on two main steps: cell sorting, and extracting the calcium transient signal over time. Most approaches [1, 4, 5, 6] mainly focus on the first step since transient signals are extracted from the raw data over time for each cell centroid detected without considering its spatial formation as shape, intensity and *etc*. Therefore, we choose to address both steps because the same idea of matrix decomposition can be applied to infer calcium transient events considering the spatial information of the detected cells. This make it easier to detect and interpret events, *e.g*. cell detection or transient waveforms, than to observe directly raw data. Specifically, the cell sorting consists of combining a sparse dictionary learning [7] to decompose raw data into a few constituent signals and an image segmentation; whereas the transient analysis is performed by modeling raw data as a linear combination of few non–zero cells extracted previously. In the following we discuss the details of each step.

2.2. Sparse Dictionary Learning for extracting cellular signals

The extraction of cellular signals is formulated as a sparse dictionary learning, thus allowing more flexibility to adapt the representation to the data, exploiting sparsity occurred in the data and not imposing that the basis vectors be orthogonal. Dictionary learning should allow to model fluorescence calcium images as a combination of few dictionary elements, each of which should be associated with a group of co–activated cells (basis functions). Specifically, we assume that the vectorized data $\mathbf{X} \in \mathbb{R}^{N\times T}$, where T is the number of frames and N is the number of pixels in an image, is decomposed into two low rank matrices: spatial component $\mathbf{D} \in \mathbb{R}^{N\times K} = [\mathbf{d}_{:1}, \ldots, \mathbf{d}_{:K}]$ and the temporal component $\mathbf{U} \in \mathbb{R}^{T\times K} = [\mathbf{u}_{:1}, \ldots, \mathbf{u}_{:K}]$; where K is the number of dictionary elements, and $\mathbf{d}_{:k}$ and $\mathbf{u}_{:k}$ represent the spatial and temporal description of the k^{th} group of cells[1]. This decomposition consists of finding the most representative basis functions and its temporal description to describe the complete data inducing sparsity over time and space. This is achieved by minimizing:

$$\min_{\mathbf{D}\in\mathcal{D},\mathbf{U}} \|\mathbf{X} - \mathbf{D}\mathbf{U}^T\|_F^2 + \lambda^U \sum_{j=1}^{K} \|\mathbf{u}_{:j}\|_1 \text{ s.t.}$$
$$\mathcal{D} \triangleq \left\{\mathbf{D} \in \mathbb{R}^{N\times K} | \forall j, \|\mathbf{d}_{:j}\|_2^2 + \gamma\|\mathbf{d}_{:j}\|_1 \leq 1\right\}, \qquad (1)$$

where λ^U is a regularization parameter that induces sparsity in the temporal coefficients; γ induces sparsity in the dictionary elements; $\|\cdot\|_F$ is the Frobenius norm; and $\mathcal{D}$ is the convex set of columns that satisfies the constraint which prevents $\mathbf{D}$ from becoming arbitrarily large. An approximate solution of Eq. (1) can be found by alternating between the two variables, minimizing over one variable while keeping the other one fixed. The details of the optimization are explained in [7]

[1] $\mathbf{d}_{:k}$ represents the k^{th} column vector of matrix $\mathbf{D}$ using MATLAB notation.

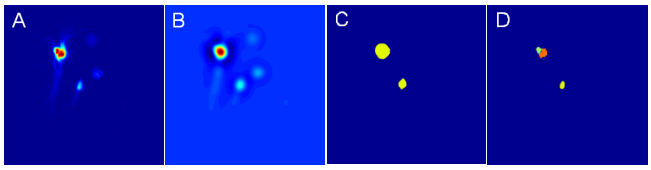

Fig. 3. Extraction of single cells. a) k^{th} basis function; b) the basis function smoothed up to the 5^{th} wavelet scale; c) events detected with non–maximal suppression and d) final definition of cell boundaries.

and to solve Eq. (1) we used the publicly available SPAMS package [7].

2.3. Image Segmentation and Transient analysis

Once basis functions are extracted, we address the problem of distinguishing overlapping and correlated cells since the sparse dictionary learning yields basis functions that may contain more than one cell as shown in Fig. 3A: overlapping cells or those with correlated activity typically are associated with the same basis function. The identification of candidates of individual cells is optimized for all $\mathbf{d}_{:k}$ in $\mathbf{D}$ by three steps as follows (Fig. 3):

1. The standard deviation (SD) σ_k of only those pixels whose intensities do not belong to calcium event is estimated for each basis $\mathbf{d}_{:k}$;

2. a wavelet transform proposed by [1] is computed up to level 5 to enhance those regions/pixels that belong to neurons above the background noise;

3. each basis k is segmented by a non–maximal suppression given a certain noise level, *i.e.* $t \cdot \sigma_k$ with $t = 3, \ldots, 6$, followed by a watershed algorithm to extract regions belonging to individual cells; then heuristics described in [1] are applied to detect candidate cells.

Finally, all the candidate cells extracted are refined in order to identify unique individual cells since one cell could appear at different basis functions. Specifically, a $2D$-Gaussian is fitted for each region of candidate cells; instead of a Mahalanobis distance, the similarity for determining whether two candidate cells describe the same active cell or not is equal to the area of the product between their fitted Gaussians. This similarity takes both distributions into account as follows:

$$S(j,q) = \frac{\exp\left(-\frac{1}{2}(\mu_j - \mu_q)^T(\mathbf{\Sigma}_j + \mathbf{\Sigma}_q)^{-1}(\mu_j - \mu_q)\right)}{2\pi\sqrt{|\mathbf{\Sigma}_j + \mathbf{\Sigma}_q|}}, \tag{2}$$

where μ_j and μ_q is the centroid of the j^{th} and q^{th} candidate cell, respectively, and Σ_j and Σ_q is the full covariance matrix of the j^{th} and q^{th} candidate cell. Hence, two candidates are fused to describe a single candidate cell if this similarity is greater than a threshold. This process runs until no more changes occurre, thus obtaining a new basis of individual cells denoted as $\mathbf{D}^S$.

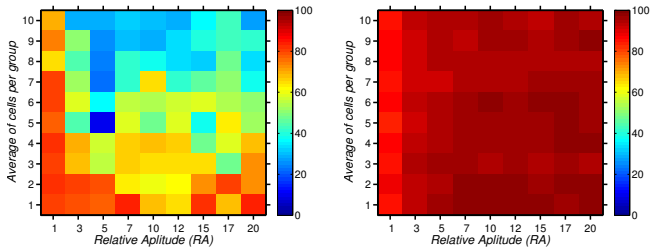

Fig. 4. Results of the algorithm sensitivity (%) for Cell Sorting [5] (left) and our approach(right).

Instead of extracting the intensities of the raw data over time at the centroid of the individual cells or considering $\mathbf{U}$, we address the estimation of the temporal component to reconstruct the data X in Eq. (1) keeping fixed the basis functions of single cells $\mathbf{D}^S$. This allows to again exploit sparsity in time and the spatial information of the individual cells $\mathbf{D}^S$ by minimizing:

$$\min_{\mathbf{U}^S} \|\mathbf{X} - \mathbf{D}^S[\mathbf{U}^S]^T\|_F^2 + \lambda^S \sum_{t=1}^{T} \|\mathbf{u}_{t:}^S\|_1 \tag{3}$$

where $\mathbf{u}_{t:}^S$ is the temporal coefficients of the individual cells at the t^{th} frame and λ^S is the regularization parameter which, for large values, makes sure that only few cells are used in the reconstruction of the t^{th} frame.

3. RESULTS

Artifical Sequences: For evaluation, we created artificial sequences with 450 frames of size 512×512 pixels. The data is created by randomly selecting cell shapes from 36 different active cells extracted from real data, and locating them into different locations with an overlap of up to 30%. Each cell is randomly assigned to a single or multiple groups of neurons that fire approximately at the same time. The existence, distribution, and activity patterns of each cell is known. In order to quantify the algorithm performance, we compute the sensitivity used in [1] that is the ratio of correctly detecting events to the number of all embedded events. We always use five groups, but vary the average number of cells per group between 1 to 10. In addition, we add random Gaussian noise σ_{noise} with a relative amplitude, $RA = \frac{I_{max} - I_{mean}}{\sigma_{noise}}$, between 1 and 20, where I_{max} and I_{mean} are the maximum and mean intensity amplitude, respectively.

As shown in Fig. 4, our method achieves a sensitivity of 94.3% on average and discriminates robustly individual cells for different levels of correlated cell activity and of noise. The best sensitivity achieved by Cell Sorting [5] is 83.1% for different noise levels, but its performance is reduced when the average number of cells per group increases.

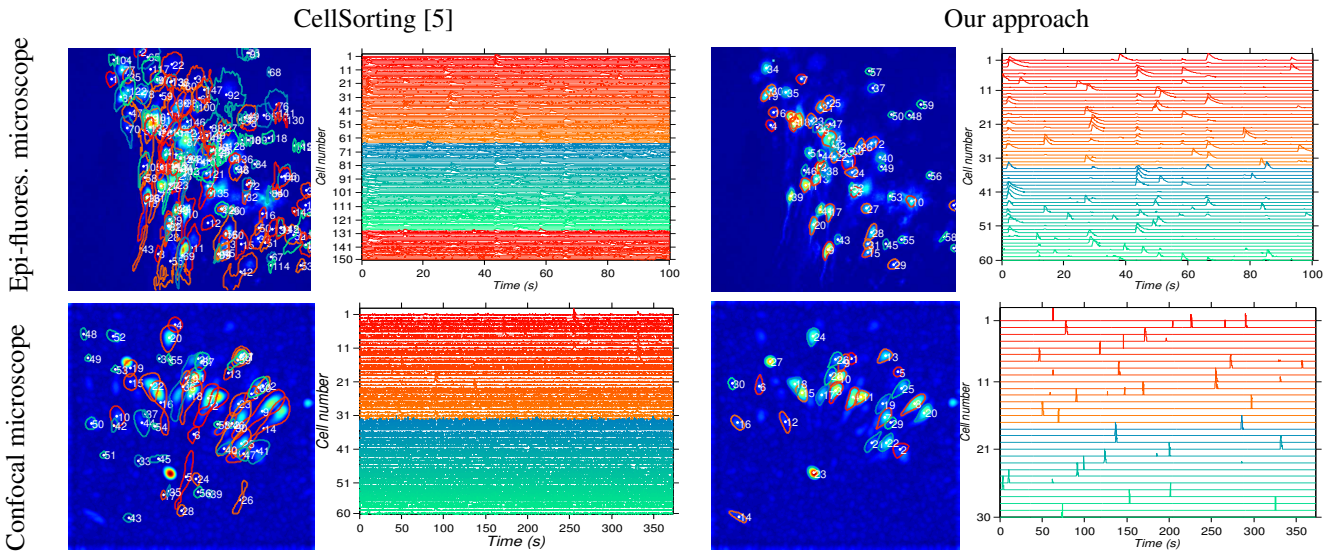

Fig. 5. Cell sorting and transient analyis using Cell Sorting [5] (left) and our approach (right) for different real sequences recorded by epi–fluorescence microscopy (top) and confocal microscopy (bottom).

Real Sequences: We applied the automated pipeline to epifluorescent and confocal calcium imaging data sets from different mice (C57BL6) hippocampal slice cultures that lacks ground truth due to the challenge of annotating neurons. As shown in Fig. 5, our method is able to distinguish overlapping cells and highly correlated cells, and also extracts smooth and regularly shaped cells. Also, as shown in Fig. 5, exploiting spatio–temporal sparsity and the shape cell information allows our approach to infer accurately the transient events, thus making it easier for interpreting visually when cells spike or not than inspecting directly raw data. Hence, we have confirmed that our method is able to detect automatically, and differentiate between, overlapping and highly correlated cells on different calcium imaging data. This is an improvement over [5] which does not work as well on epi fluorescent data.

4. CONCLUSIONS & DISCUSSION

We have presented an automated pipeline for identifying cell activity from calcium imaging data, and reported its performance on artificial and real sequences. In contrast to previous work, our approach formulates cell sorting and transient analysis in the same framework. This formulation allows to exploit sparsity both in the spatial and the temporal domain, and works with confocal and even epi–fluorescent raw data (which is easier to record and is readily implemented in in-vivo studies).

In the future, we plan to improve the pipeline as: 1) improve sparse dictionary learning for only extracting single cells; 2) enhance robustness of detecting calcium transients against background signals. In addition, we plan to set individual thresholds for individual cells for correlating the rate and the intensity of action potentials in order to detect which kind of cell shows this calcium transient.

5. REFERENCES

[1] S. Reichinnek, A. von Kameke, A. M. Hagenston, E. Freitag, F. C. Roth, H. Bading, M. T. Hasan, A. Draguhn, and M. Both, "Reliable optical detection of coherent neuronal activity in fast oscillating networks in vitro," *NeuroImage*, vol. 60, no. 1, 2012.

[2] X. Li, G. Ouyang, A. Usami, Y. Ikegaya, and A. Sik, "Scale-free topology of the ca3 hippocampal network: A novel method to analyze functional neuronal assemblies," *Biophysical Journal*, vol. 98, no. 9, pp. 1733 – 1741, 2010.

[3] I. Ozden, H. M. Lee, M. R. Sullivan, and S. S. Wang, "Identification and clustering of event patterns from in vivo multiphoton optical recordings of neuronal ensembles," *J Neurophysiol*, 2008.

[4] I. Valmianski, A. Y. Shih, J. D. Driscoll, D. W. Matthews, Y. Freund, and D. Kleinfeld, "Automatic identification of fluorescently labeled brain cells for rapid functional imaging," *Journal of Neurophysiology*, 2010.

[5] E. A. Mukamel, A. Nimmerjahn, and M. J. Schnitzer, "Automated analysis of cellular signals from large-scale calcium imaging data," *Neuron*, 2009.

[6] R. Maruyama, M. Kazuma, M. Hiroyoshi, and A. Toru, "Detection of cells from calcium imaging data using non-negative matrix factorization," in *Proceedings of the 21st Annual Conference of the Japanese Neural Network Society*, 2011.

[7] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proceedings of the 26th Annual International Conference on Machine Learning*, 2009, ICML '09.