HOTARU: Automatic sorting system for large scale calcium imaging data

Takashi Takekawa[1,2,*], **Masanori Nomoto**[3,4,†], **Hirotaka Asai**[3,4,†], **Noriaki Ohkawa**[5,6], **Reiko Okubo-Suzuki**[3,4], **Khaled Ghandour**[3,4,7], **Masaaki Sato**[2,8,9,10], **Masamichi Ohkura**[9,10], **Junichi Nakai**[11], **Shin-ichi Muramatsu**[12,13], **Yasunori Hayashi**[2,10,14], **Kaoru Inokuchi**[3,4], **Tomoki Fukai**[2,15]

***For correspondence:**
takekawa@cc.kogakuin.ac.jp (TT)

†These authors contributed equally to this work

[1]Faculty of Informatics, Kogakuin University of Technology and Engineering, Tokyo, Japan; [2]RIKEN Brain Science Institute, Saitama, Japan; [3]Graduate School of Medicine and Pharmaceutical Sciences, University of Toyama, Toyama, Japan; [4]Research Centre for Idling Brain Science, University of Toyama, Toyama, Japan; [5]Comprehensive Research Facilities for Advanced Medical Science, Dokkyo Medical Univerity, Tochigi, Japan; [6]PRESTO, Japan Science and Technology Agency (JST), Saitama, Japan; [7]Department of Biochemistry, Faculty of Pharmacy, Cairo University, Egypt; [8]Department of Neuropharmacology, Hokkaido University Graduate School of Medicine, Sapporo, Japan; [9]Graduate School of Science and Engineering, Saitama University, Saitama, Japan; [10]Brain and Body System Science Institute, Saitama University, Saitama, Japan; [11]Graduate School of Dentistry, Tohoku University, Sendai, Japan; [12]Division of Neurological Gene Therapy, Jichi Medical University, Shimotsuke, Japan; [13]Center for Gene & Cell Therapy, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [14]Graduate School of Medicine, Kyoto University, Kyoto, Japan; [15]Neural Coding and Brain Computing Unit, Okinawa Institute of Science and Technology, Okinawa, Japan

Abstract Currently, calcium imaging allows for the long-term recording of large-scale neuronal activity in diverse states. However, it remains difficult to extract neuronal dynamics from recorded imaging data. In this study, we propose an improved CNMF-based algorithm and an effective method for extracting cell shapes with fewer false positives and false negatives caused by image processing. We also showed that the values obtained during image processing can be combined and used for false positive determination of cells. For the CNMF algorithm, we combined cell-by-cell regularization and baseline shrinkage estimation, which greatly improved its stability and robustness. We applied these methods to artificial and real data and confirmed their effectiveness. Our method is simpler and faster, detects more cells with lower firing rates and signal-to-noise ratios, and enhance the quality of the extracted cell signals. These advances can improve the standard of downstream analysis and contribute to progress in neuroscience.

Introduction

A major goal of neuroscience is to understand interactions within large populations of neurons, including network dynamics and emergent behaviors. This requires recording diverse activities in large populations of neurons. In recent years, genetic indicators have been widely recorded by microscopy, and advances in genetic optics have greatly improved the signal strength of calcium-

binding protein indicators (***Chen et al., 2013***). Currently, neuronal activity can be imaged in vivo in a wider region of the brain for longer periods (***Cotton et al., 2013***; ***Ahrens et al., 2013***; ***Prevedel et al., 2014***; ***Diego Andilla and Hamprecht, 2013***). In addition to two-photon microscopy, the development of miniaturized microscopes such as microendoscopes has greatly expanded the area that can be visualized, such as the deep brain of a freely moving mouse, and the data obtained has burgeoned (***Flusberg et al., 2008***; ***Ghosh et al., 2011***; ***Ziv and Ghosh, 2015***). Therefore, an efficient and accurate method is required to detect the spatial positions and shapes of neurons as well as their temporal firing patterns from the obtained imaging data.

The first method for extracting individual cells from calcium imaging data is region of interest (ROI) analysis, which first extracts an ROI and then extracts the temporal changes in calcium concentration within the ROI (***Smith and Häusser, 2010***; ***Pinto and Dan, 2015***; ***Barbera et al., 2016***; ***Klaus et al., 2017***). However, this method cannot effectively separate the signals of spatially overlapping neurons. In addition, there is no accurate and efficient method for obtaining ROIs for large data sets (***Resendez et al., 2016***). The second method is PCA / ICA analysis (***Mukamel et al., 2009***). Because this method is a linear separation method, it still cannot accurately separate cells with respect to spatial overlap (***Resendez et al., 2016***). Consequently, a third method, that includes non-negative matrix factorization (NMF) (***Maruyama et al., 2014***) and constrained NMF (CNMF) analysis (***Pnevmatikakis et al., 2013***, ***2016***), has been proposed. CNMF analysis models the position and shape of cells (footprint), the relationship between spikes and calcium fluctuations, and the statistical nature of noise. Therefore, it is possible to deconvolve overlapping cells into pure activity without causing mutual noise contamination. However, algorithms that simply optimize the model are computationally expensive. The standard solution uses an iterative algorithm that alternately improves calcium fluctuations and cell shapes. But instabilities arise during this iteration that degrade accuracy. For this reason, various computational innovations have been proposed. Among them, CNMF-E is a widely used, high-performance method that considers the effects of background fluctuations in addition to numerous computational innovations (***Zhou et al., 2018***).

However, CNMF-E also presents some areas for improvement. It uses three parameters for initialization: cell size, intensity (signal-to-noise ratio), and a threshold for correlation with neighboring pixels. The greedy method detects cells by order of signal size, a rather complicated and arbitrary procedure that repeats a small algorithm that analyzes the spatio-temporal structure of single cells at each step. Particularly for regions such as the hippocampus, where cell size variation is large, it is difficult to set parameters, requiring multiple attempts to adjust them, including by manual procedures.

Another drawback is that the iterative method computes the background in a step independent of the cellular activity. The CNMF-based algorithm does not penalize the background activity for footprint and fluorescence intensity variations, variables that are penalized due to sparsity limitations arising from its structure. Therefore, if background activity is regarded as a component, the risk is that fluctuations that are essentially cellular activity will be forced into the background activity. Otherwise expressed, the determination of the reference baseline is critical for estimating the non-negative cellular activity, and since cellular activity and the baseline significantly influence each other, it is a difficult question whether the iterative method is sufficiently effective. Specifically, the same fluorescence intensity fluctuation can be estimated as a continuous firing activity or a low-frequency firing state, depending on how the baseline is taken (***Takekawa et al., 2017***).

In this paper, we propose HOTARU (High performance Optimizer to extract spike Timing And cell location from calcium imaging data via lineaR impUlse), a simple, automatic calcium imaging data analysis system that improves on the above two limitations. First, multiple cell size parameters are set within a certain range, and an initial footprint is formed using a Laplacian of Gaussian (LoG) filter and watershed cuts algorithm (***Cousty et al., 2008***), which is effective for blob detection (***Lindeberg, 2013***). In addition, candidate cells are detected independently from every frame of the video to prevent missing those with low firing rates. This means that a huge number of cell candidates are generated in the first step, from which a greedy method that takes advantage of

Table 1. Variables and parameters used in the HOTARU model.

kind	parameter	description
Data	$\mathrm{F}[t,x]$	normalized video data
	T	number of frames
	X	number of pixels
Independent parameters	$\gamma[\tau]$	response function ($\Gamma[t-\tau,t]=\gamma[\tau]$)
	ξ_{T}	coefficient of temporal baseline penalty
	ξ_{X}	coefficient of spatial baseline penalty
	λ_{U}	coefficient of spike penalty
	λ_{A}	coefficient of footprint penalty
Independent variables	K	number of components
	$a_k[x]$	footprints
	$u_k[t]$	spike trains
Dependent parameters	ξ'_{T}	shrinkage parameter $\left(=\left(1+\xi_{\mathrm{X}}\right)^{-1}\right)$
	ξ'_{X}	shrinkage parameter $\left(=\left(1+\xi_{\mathrm{T}}\right)^{-1}\right)$
Operators	$E_t y[t]$	Average operator $\frac{1}{T}\sum_t y[t]$
	$\boldsymbol{B}_{\xi}$	Baseline operator $\xi'_T\left(E_t-E_{t,x}\right)+\xi'_X\left(E_x-E_{t,x}\right)+E_{t,x}$
Dependent variables	$v_k[t]$	calcium fluctuation $\left(=\sum_\tau u_k[t-\tau]\gamma[\tau]\right)$
	$b[t,x]$	baseline $\left(=\boldsymbol{B}_{\xi}\left(\mathrm{F}[t,x]-\sum_k a_k[x]v_k[t]\right)\right)$

the size of cells with variations, efficiently redues the number of candidate cells to the minimum necessary. There are no arbitrary parameters in the algorithm, and it operates automatically.

HOTARU also makes some modifications to the basic CNMF model and the optimization algorithm is proposed accordingly. L1 regularization is performed for normalized footprints and spike firings per footprint. By using a proximal operator for group-normalized L1 regularization, we can define an algorithm that extends the proximal gradient method with Nesterov's (***Nesterov, 1983***, ***1988***). The background variation is estimated simultaneously with cellular activity by assuming that background variation is independent in time and space, and that background variation is the dependent variable for footprint and fluorescence intensity changes. Furthermore, by setting a normally distributed prior on the variance of the background variation, we avoid indeterminacy in the baseline estimation by penalizing it with a reduced estimation of the background variation. Mathematically, this prior corresponds to a change in the diagonal components of the matrix used in the optimization problem and affords increasing computational stability.

Model and model fitting

Basic concept

The recoded video data can be represented by the matrix $F[t,x]\in\mathbb{R}^{T\times X}$, where T is the number of frames and X is the number of pixels in the target area. In our's, as in the CNMF model the activity of each neuro k is modeled by multiplying a spatial footprint vector $a_k[x]\in\mathbb{R}_+^X$ representing the location and shape of the cell and calcium activity vector $v_k[t]\in\mathbb{R}_+^T$. Both $a_k[t]$ and $v_k[t]$ are restricted to be non-negative by the physical interpretation. However, since indeterminacy arises for $a_k[x]$ and $v_k[t]$, we restrict them to satisfy the condition $\max_x a_k[x]=1$. If the recording region contains K cells, the recorded video data is modeled by the sum of the spatio-temporal activity of all cells, background variation $\varepsilon[t,x]\in\mathbb{R}^{T\times X}$ which incluses additive noise:

$$F[t,x]=\sum_{k=1}^{K}a_k[x]v_k[t]+\varepsilon[t,x]. \tag{1}$$

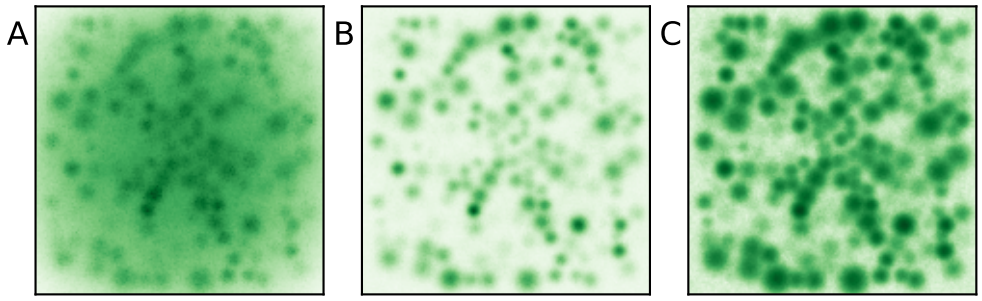

Figure 1. Simulation video: the maximum (A), standard deviation (B), and local correlation (C).

Calcium fluctuation $v_k[t]$ is represented by the linear sum of response function $\gamma[\tau] \in \mathbb{R}_+$ to spikes $u_k[t] \in [0, 1]$:

$$v_k[t] = s_k \sum_{\tau=1}^{\tau_{\max}} \gamma[\tau] u_k[t-\tau]. \tag{2}$$

where $s_k \in \mathbb{R}_+$ corresponds to cell k's fluorescence intensity. Although any response function can be used as a model and a solution method, this paper presents the results of using a general second-order AR model and the corresponding double exponential function.

Baseline regurization

The video is assumed to be subject to background variations corresponding to time $b_T[t] \in \mathbb{R}^T$ and space $b_T[t] \in \mathbb{R}^T$, respectively:

$$\varepsilon[t, x] = b_T[t] + b_X[x] + b_0[t, x]. \tag{3}$$

It is assumed that b_T, b_X, and b_0 obey normal distributions. In this case, the mean of b_T and b_X does not lose generality even if the mean of b_T and b_X is 0. The standard deviations of b_T and b_X were assumed to be the constant times the standard deviation of b_0:

$$b_0[t, x] \sim \mathcal{N}\left(\mu, \sigma^2\right), \quad b_T[t] \sim \mathcal{N}\left(0, \zeta_T^{-1}\sigma^2\right), \quad b_X[x] \sim \mathcal{N}\left(0, \zeta_X^{-1}\sigma^2\right). \tag{4}$$

Minimizing the variance of the whole $b[t, x]$ under these assumptions, $b[t, x]$ can be expressed in a dependent manner by letting $F - AV$ act on the operator B_ζ (see ***Table 1***).

Normalized group L1 regurization

Because cells are spatially localized, footprints $a_k[x]$ have the sparse property of being 0 in many locations. The local structure is not explicitly incorporated into the model, but sparsity is represented by the L1 regularization term. However, the restriction of $\max_x a_k[x] = 1$ requires a special solution method. Similarly, for spike $u_k[t]$, we assume the sparsity of the firing rate rather than the change in fluorescence intensity itself under the condition $u_k[t] \leq 1$. As such, a normalized L1 regularization term is introduced as in $a_k[x]$. Given the assumption of MAP inference, the addition of the L1 regularization term corresponds to assuming a Laplace distribution for the variables. The Laplace distribution is idential to an exponential distribution $\mathcal{E}$ when the variables are non-negative:

$$a_k[x] \sim \mathcal{E}\left(\lambda_A\right), \quad u_k[t] \sim \mathcal{E}\left(\lambda_U\right). \tag{5}$$

Further details are provided in the Materials and methods section.

Fitting algorithm

The MAP inferece of model ***Equation 1***, ***Equation 2*** and the prior ***Equation 4***, ***Equation 5*** maximizes the following logarithm of posterior likelihood L:

$$L = -\frac{TX}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t,x}\left(F[t,x] - \sum_k s_k a_k[x]\sum_\tau \gamma[\tau - t]u_k[t] - b[t,x]\right)^2 - \frac{X}{2}\log\sigma^2 - \frac{T\zeta_X}{2\sigma^2}\sum_x b_X[x] - \frac{T}{2}\log\sigma^2 - \frac{X\zeta_T}{2\sigma^2}\sum_t b_T[t] - \lambda_A\sum_{k,x} a_k[x] - \lambda_U\sum_{k,t} u_k[t]. \quad (6)$$

The following conditions must be satisfied:

$$s_k \geq 0, \quad a_k[x] \geq 0, \quad u_k[t] \geq 0, \quad \max_x a_k[x] = 1, \quad \max_t u_k[x] = 1. \quad (7)$$

Here, we denote the parameters in the matrix form by $F \in \mathbb{R}^{X\times T}$, $A \in \mathbb{R}_+^{K\times X}$, and $A \in \mathbb{R}_+^{K\times X}$, respectively. Even if the data is centralized so that $\sum_x F = 0$ and $\sum_t F = 0$, the generality of the solution is preserved and $B_\zeta F = 0$ always holds for the operation B_ζ. Considering the maximization of L regarding σ^2, the solution of ***Equation 6*** satisfies the following condition:

$$E = (TX + T + X)\,\sigma^2 = \left\|F - A^\top V + B_\zeta A^\top V\right\|_2^2, \quad (8)$$

where $\|\cdot\|_2$ is the Frobenius norm. Thus, we can simplify the problem as follows:

$$\hat{L}[s,a,u] = \log E + \hat{\lambda}_A\sum_{k,x} a_k[x] + \hat{\lambda}_U\sum_{k,t} u_k[t], \quad (9)$$

Then, by defining $\hat{a}_k[x] = s_k a_k[x]$ and $\hat{u}_k[t] = s_k u_k[t]$, we can split this problem into the following two sub-problems:

$$\underset{\hat{A}}{\text{maximize}} \quad \log\left\|F - V_*\hat{A} + B_\zeta V_*\hat{A}\right\|_2^2 + \sum_k \frac{\hat{\lambda}_A\sum_x \hat{a}_k[x]}{\max_x \hat{a}_k[x]} \quad \text{subject to} \quad \hat{a}_k[x] \geq 0, \quad (10)$$

$$\underset{\hat{U}}{\text{maximize}} \quad \log\left\|F - A_*\Gamma\hat{U} + B_\zeta A_*\Gamma\hat{U}\right\|_2^2 + \sum_k \frac{\hat{\lambda}_U\sum_t \hat{u}_k[t]}{\max_t \hat{u}_k[t]} \quad \text{subject to} \quad \hat{u}_k[t] \geq 0. \quad (11)$$

Both ***Equation 10*** and ***Equation 11*** consist of a logarithm of quadratic form and a regularized L1 regularization term, which can be generalized as follows:

$$L[X] = \log g[X] + \sum_i \frac{\lambda\sum_j x_{ij}}{\max_j x_{ij}}. \quad (12)$$

The proximal operator for this regularization term can be constructed as follows:

$$\text{prox}_L[x^*] = \max\left\{0, x^* - \eta\left(\frac{\nabla g[X]}{g[X]} + \xi_{ij}\right)\right\}, \quad \xi_{ij} = \begin{cases} 0 & x_{ij} = \max_k x_{ik} \\ \lambda/\max_k x_{ik} & x_{ij} \neq \max_k x_{ik} \end{cases}, \quad (13)$$

and the optimization problem can be solved by repeatedly applying the proximal operator. In addition, speed-ups such as Nesterov acceleration can be achieved during iterative application. In both ***Equation 10*** and ***Equation 11***, the target function g is the quadratic form of the variable vector. Therefore, by computing the necessary matrices in advance, only simple matrix products are needed at iteration. Consequently, the proposed algorithm can be processed at high-speed using GPUs' parallel operations.

Image processing

A LoG filter is a spatial filter with size parameter r used in image processing to detect edges and blobs, which is expressed as the following equation:

$$\text{LoG}_r = \frac{1}{\pi r^2}\left[\frac{\|x\|^2}{2r^2} - 1\right]\exp\left[-\frac{\|x\|^2}{2r^2}\right]. \quad (14)$$

In our method, the LoG filters with multiple size parameters $\{r\}_{i=1}^{N_r}$ are used to create an initial footprint, and to shape and evaluate the footprint obtained by an iterative method.

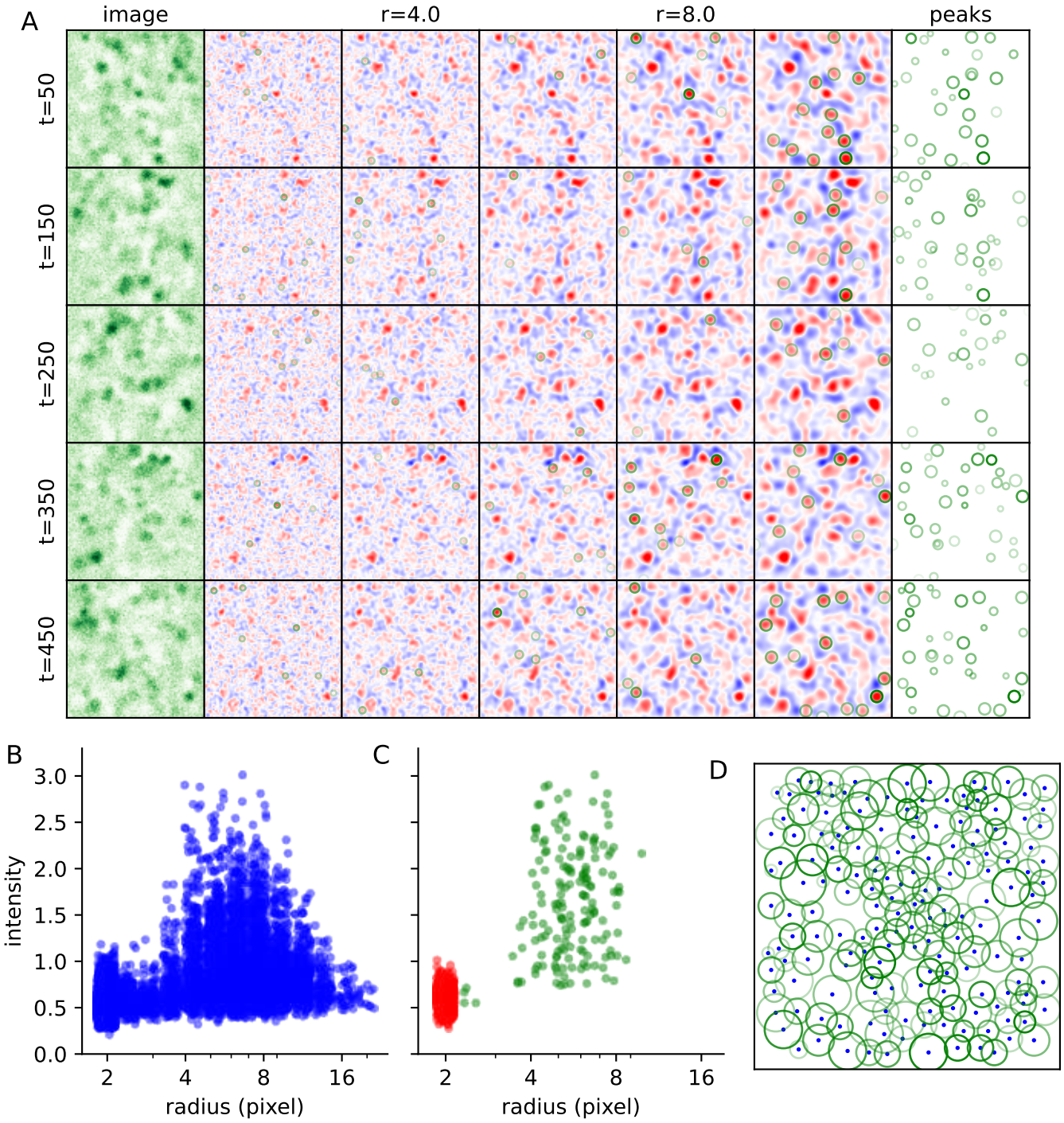

Figure 2. Detection of cell candidate peaks: (A) Example of applying a LoG filter with different r to each frame of the video to detect the maxima. (B) Distribution of filter size r and maximum value when collapsed per pixel after detection. (C) Distribution after cells are reduced using the greedy method with the cell overlap criteria $\theta = 1.2$. (D) Overall view of the obtained peaks.

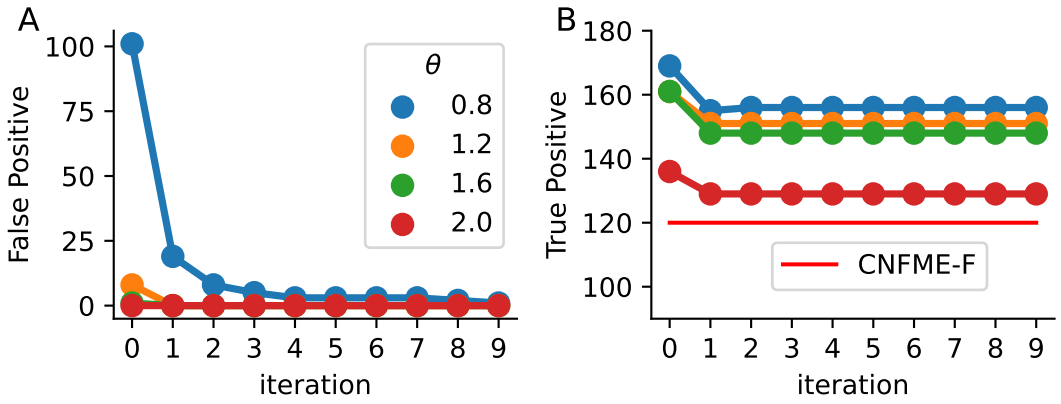

Figure 3. Effects of repeated improvement treatments: (A) changes in false positives; (B) changes in true positive.

Peak finding by LoG filter

In creating the initial footprint, each frame t of the video $F[t]$ is filtered by LoG_{r_i} of different sizes r_i are applied to the image:

$$z_t[i,x] = \mathrm{LoG}_{r_i} * F_t. \tag{15}$$

Then we collect all the local maxima (peaks) in (i,x). Next, the peaks (i,x) with the largest 'intensity' $z[i,x]$ are adopted from the large number of peaks obtained by the greedy method so that the distance between the peaks employed is always less than θr. The algorithm for efficiently obtaining a set of peaks that satisfy this condition will be described later. θ is a parameter that represents the distance between cells, otherwise expressed, the possible cell overlap.

Segmentation of footprint

The initial footprint is from the filtered image corresponding to the obtained peak, the watershed method (***Cousty et al., 2008***) is applied to the area around the peak to extranct a region in which a cell exists. For the iterative footprint, multiple LoG filters are similarly applied, the largest peak is selected, and the watershed method is used to calculate the area to be cut out. Each footprint a_k is limited to a maximum value of 1, thus the z value obtained by the LoG filters is interpreted differently than in the case of initial footprint and is referred to as 'firmness' $\hat{z}$ in our paper. The obtained values $\hat{z}$ and r at the peak are useful for determining false positives.

Full details on the algorithms for initializing and then solving these three sub-problems are provided in the Materials and methods section.

Results

HOTARU reliably detect weak and low-firing-rate cells

First, as an initial condition, artificial data were used to verify the efficacy of the proposed cell-detection approach. Cells in the artificial data varied greatly in their sizes, firing frequencies, and signal-to-noise ratios, and included some that were extremely difficult to detect (***Figure 1***). As an initial condition, a LoG filter was applied to each frame of the video (***Figure 2***A left) (***Figure 2***A middle) to obtain a pair of peak positions, intensities, and radii (***Figure 2***A right). The frame-by-frame calculation allows detection of cells with low firing rates. Small signals are not excluded at this point, leaving little chance of missing a cell.

The resulting scatterplot of radii and intensity consists of roughly two groups (***Figure 2***B). Peaks around radii of two pixels correspond to false positives, and peaks between four and eight pixels in radius correspond to true cell candidates. When these peaks are condensed using the greedy

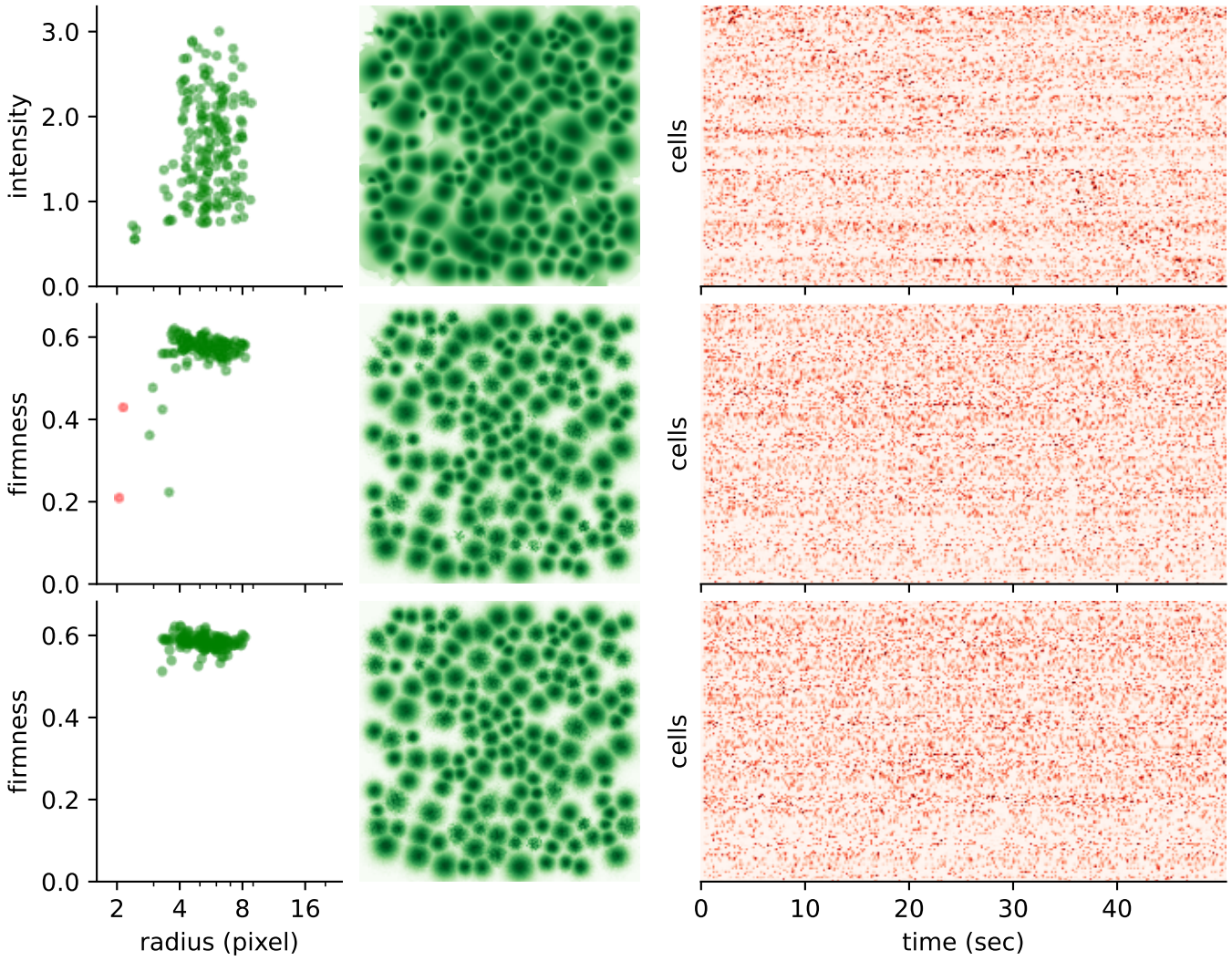

Figure 4. Effects of repeated improvement process: (A) distribution of intensity in the initial placement; (B) distribution of firmness after one application of the improvement process; (C) distribution of firmness finally obtained; (D-F) overlaid images of the footprints at each stage; (G-I) obtained spike time series.

Figure 4–video 1. Video comparing CNMF-E and HOTARU results for artificial data.

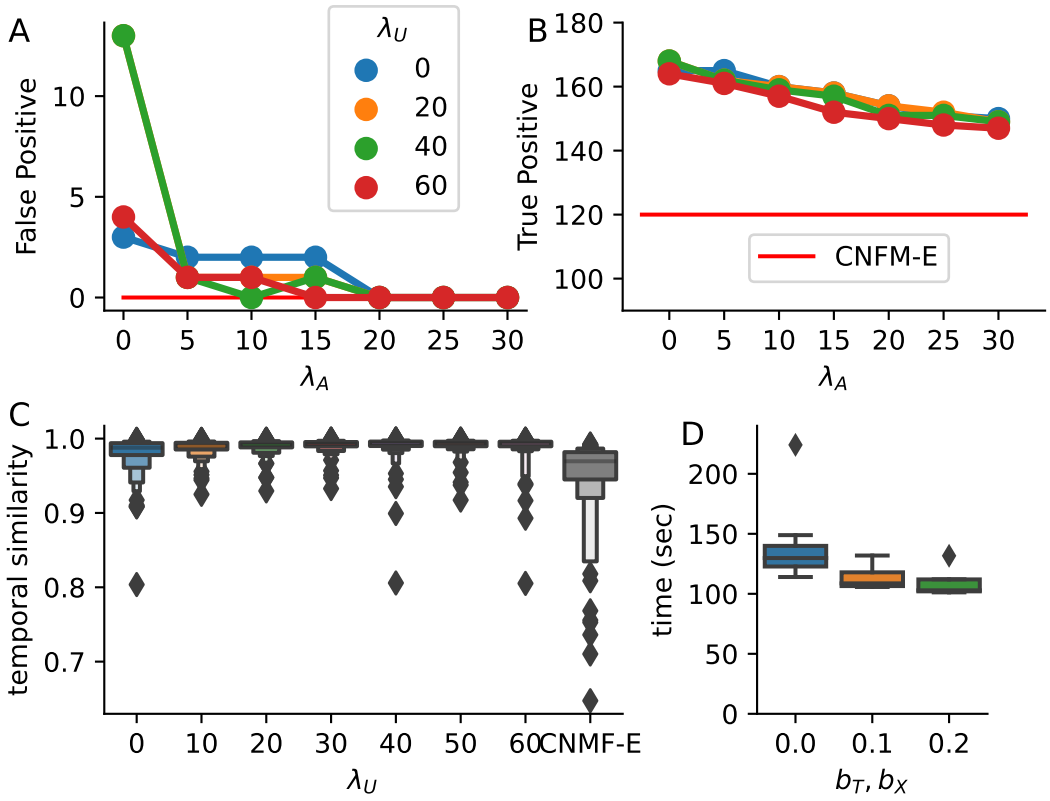

Figure 5. (A) false positives per parameter λ_U and λ_A; (B) number of detected cells per parameter λ_U and λ_A; (C) accuracy of time series estimation per λ_U; (D) variation of computation time with b.

method with cell radius information, the two groups are more clearly separated (***Figure 2***C). Eliminating the peaks corresponding to the twp-pixel radius filter as false positives allows for accurate cell detection.

Figure 2D shows the position, radii, and intensities of the obtained cells simultaneously. By using multiple filters with different radii, the distribution of peak intensities of potentially false positive cells overlaps with the distribution of peak intensities of actual cells, thus enabling a reliable false positive determination. Because the only parameters to be set are those that control the range of the cell radius and the allowed overlap of cells, whether the parameters are appropriate is easy to read from the results in ***Figure 2***C and is a major practical advantage for reducing analytical arbitrariness.

HOTARU consistently eliminates false positives

The number of cells detected depends on the parameters that controll cell overlap. If this parameter is small, a single cell may be split into multiple candidates or contain false positives that are difficult to distinguish. Conversely, if the value of this parameter is too large, one of the adjacent cells may be missed. However, our algorithm intends that such initial state irregularities would be ameliorated by iterations of the optimization algorithm.

Indeed, the number of false positives becomes zero after repeated iterations, even if the initial condition contains a large number of false positives (***Figure 5***). However, the number of correctly detected cells remains constant (***Figure 5***). The start has more candidate cells and more false positives and allows a higher final number of cells to be detected. It also inccreases the required number of iterations, indicating a trade-off. Ultimately, HOTARU correctly detected about 150 cells on average and more than 125 cells even with inappropriate parameters, while 120 cells were

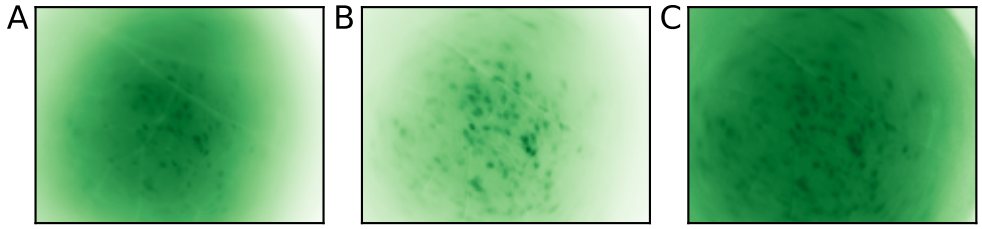

Figure 6. Data 1, hippocampal CA1 data by miniature microscopy: the maximum (A), standard deviation (B), and local correlation (C).

detected with the standard CNMF-E parameters. Neigher method detected false positives for this data.

HOTARU improves cell shape

Next, we examine in detail the elimination of false positives and the quality of the estimation results by iteration. HOTARU starts with an initial footprint and iterates between estimaing spike firing and background activity or footprint and background activity, and post-processing to mask cellular regions. The MAP inference algorithm, which includes a prior distribution, reduces the number of uncertain cell candidates as an effect of automatic relevance determination (ARD). During post-processing, the cell radius is calculated and cells are removed if the cell radius falls outside the set range; like in CNMF-E, interventions such as merging, splitting, deleting, or adding undetected cells are not performed.

For the initial footprint (***Figure 4***), the cell radii of some (red) cell candidates are reduced and removed after the first iteration. In this case, most of the remaining cells have a firmness of approximately 0.55, but some cells have a smaller value. After further iterations, the firmness of most of the cells eventually clusters around 0.55, and there are no cells with firmness in the range of 0.0–0.3. This result indirectly indicates that actual and not arbitrary cell candidates were detected.

HOTARU correctly separates overlapping cell activities and is stable to parameter changes

The parameters of the HOTARU iteration are footprints, spikes, and the respective normalization terms for background variation in time and space. The cell-radius range used in cell detection was used for post-processing and firmness calculations. $\theta = 1.2$ was used to test the extent to which the four normalization terms changed the results. If the normalization term for footprints is zero and the value of λ_U is small, false positive cells cannot be removed (***Figure 5***B). In contrast, when a is greater than 20, false positives are zero regardless of the value of λ_U. In this case, the number of correctly detected cells tends to decrease relateive to λ_A. However, even when λ_A is 30, the number of correctly detected cells is 130, which is higher than 120 in CNFM-E (***Figure 5***A).

The similarity was calculated between the estimated spike activity and the resulting de-noised fluorescence intensity variation and the true fluorescence intensity variation was calculated. On average, the similarity was higher than that of CNMF-E, regardless of the parameter values (***Figure 5***C). This indicates that the activity is correctly separated for overlapping cells, with $\lambda_U > 0$ being more stable and similar for all cells compared to $\lambda_U = 0$. The best parameter for this data would be from 30–40, with some low similarity candidates at $\lambda_U = 40$, corresponding to cells with low signal strength and low firing rates not detected by the other parameters.

For $b_{\mathtt{space}}$ and $b_{\mathtt{tempo}}$, the larger b, the faster the convergence of the optimization steps of the iterations and the effect of a shortened total computation time was observed (***Figure 5***D). For b greater than 0.3, the number of detected cells decreased, while in the 0.0–0.2 range and there was no significant difference in the results.

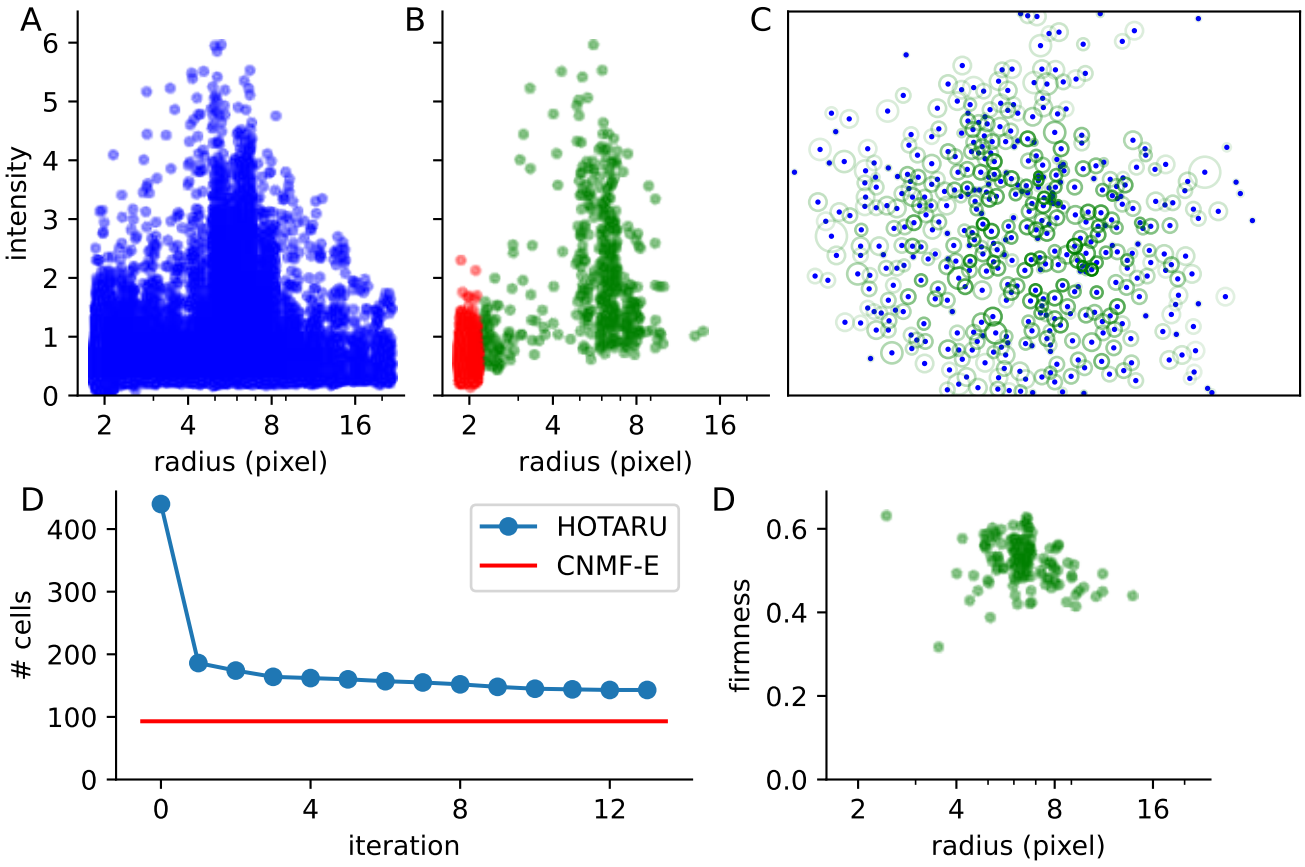

Figure 7. Summary of data 1: (A) distribution of filter size r and maximum value when collapsed per pixel after detection; (B) distribution after cells are reduced using the greedy method with the cell overlap criteria $\theta = 1.6$; (C) Overall view of the obtained peaks. (D) number of estimated cells after repeated improvement treatments. (E) distribution of cell radius and firmness at final step.

Application of HOTARU to the data from hippocampal CA1 obtained by miniature microscopy

Below is an analysis of data recorded by microendoscopy from mice (***Figure 6***) (Full experimental details and algorithm parameter settings for this and the next dataset are given in the Methods and Materials section). The video consists of 1200 frames of 720×540 pixel images taken at 20 Hz for 60 sec. We applied both HOTARU and CNMF-E to the data set and compared their performance. During initial footprint detection, the peaks of all the frames are similar to the distribution of the artificial data (***Figure 7***ABC). The maximum radius size of LoG filter $r_{\max}$ and the neuron overlap criteria θ was set to 20.0 and 1.6, respectively. and The distribution after the reduction step shows a slightly larger number of false positive cell candidates near the two-pixel cell radius compared to the artificial data. In this case, the cell candidates are clustered in the center where cells are concentrated, and no cell candidates are detected from the surrounding area. It is important to note that these results can be obtained without setting a specific threshold. After repeated iterations, the final footprint distribution of cell candidates was neatly distributed around seven pixels and firmness was 0.5.

HOTRU extracted 143 cells from this dataset while CNMF-E detected 92 cells (***Figure 7***D, ***Figure 8***A). The 84 cells detected by both CNMF-E and HOTARU exhibited almost identical behavior regarding calcium fluctuations (***Figure 8***D), while the 59 cells detected by HOTARU only showed no irregularities in shape or calcium response (***Figure 8***CD).

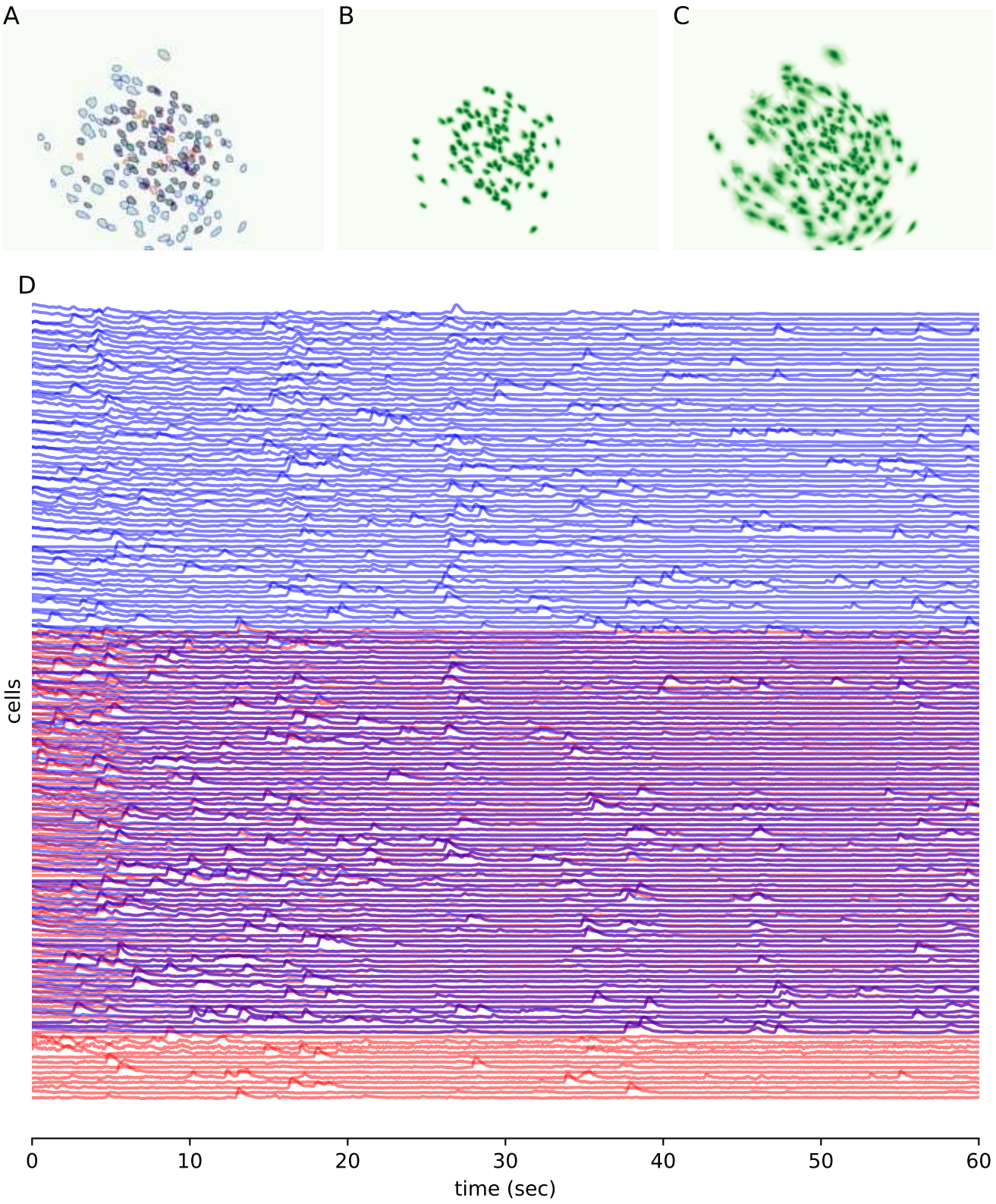

Figure 8. Comparison of CNMF-E and HOTARU on data 1: (A) CNMF-E (red) and HOTARU (blue) footprint contours superimposed; (B) footprint obtained with CNMF-E; (C) footprint obtained with HOTARU; (D) estimated cell-by-cell calcium variability by CNMF-E (red) and HOTARU (blue). Overlapping areas indicate that the footprints of the candidate cells are similar.

Figure 8–video 1. Video comparing CNMF-E and HOTARU results for data 1.

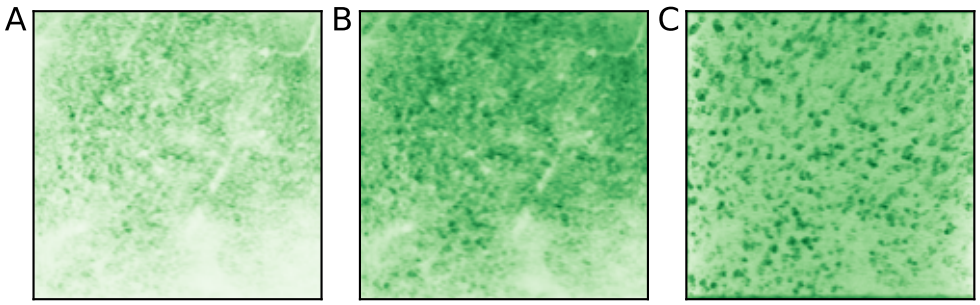

Figure 9. Data 2, the hippocampal data by two-photon microscopy: the maximum (A), standard deviation (B), and local correlation (C).

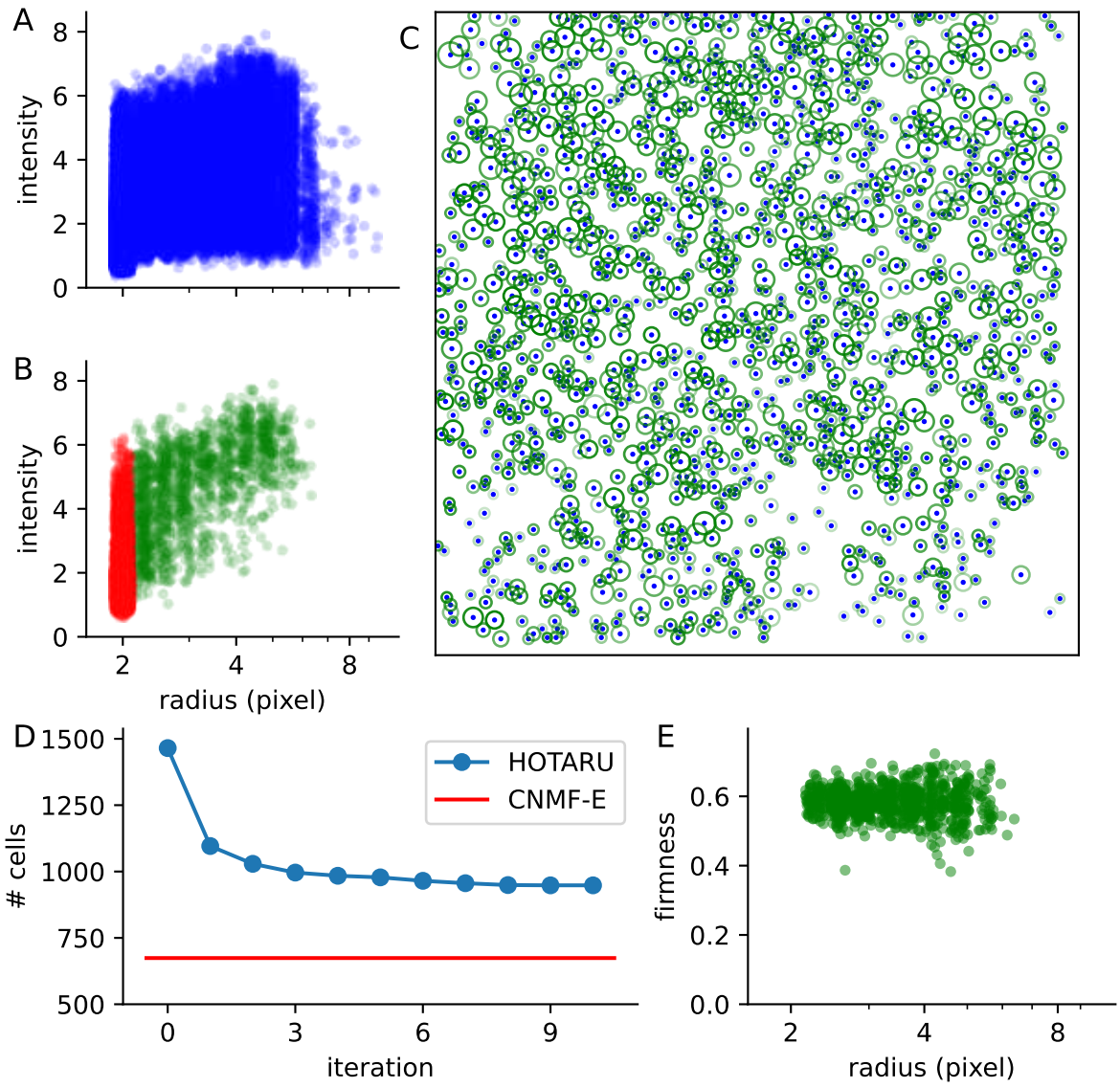

Figure 10. Summary of data 2: (A) distribution of filter size r and maximum value when collapsed per pixel after detection; (B) distribution after cells are reduced using the greedy method with the cell overlap criteria $\theta = 1.6$; (C) Overall view of the obtained peaks. (D) number of estimated cells after repeated improvement treatments. (E) distribution of cell radius and firmness at final step.

Figure 10–Figure supplement 1. Detailed distribution of radius and firmness and activity intensity s_k.

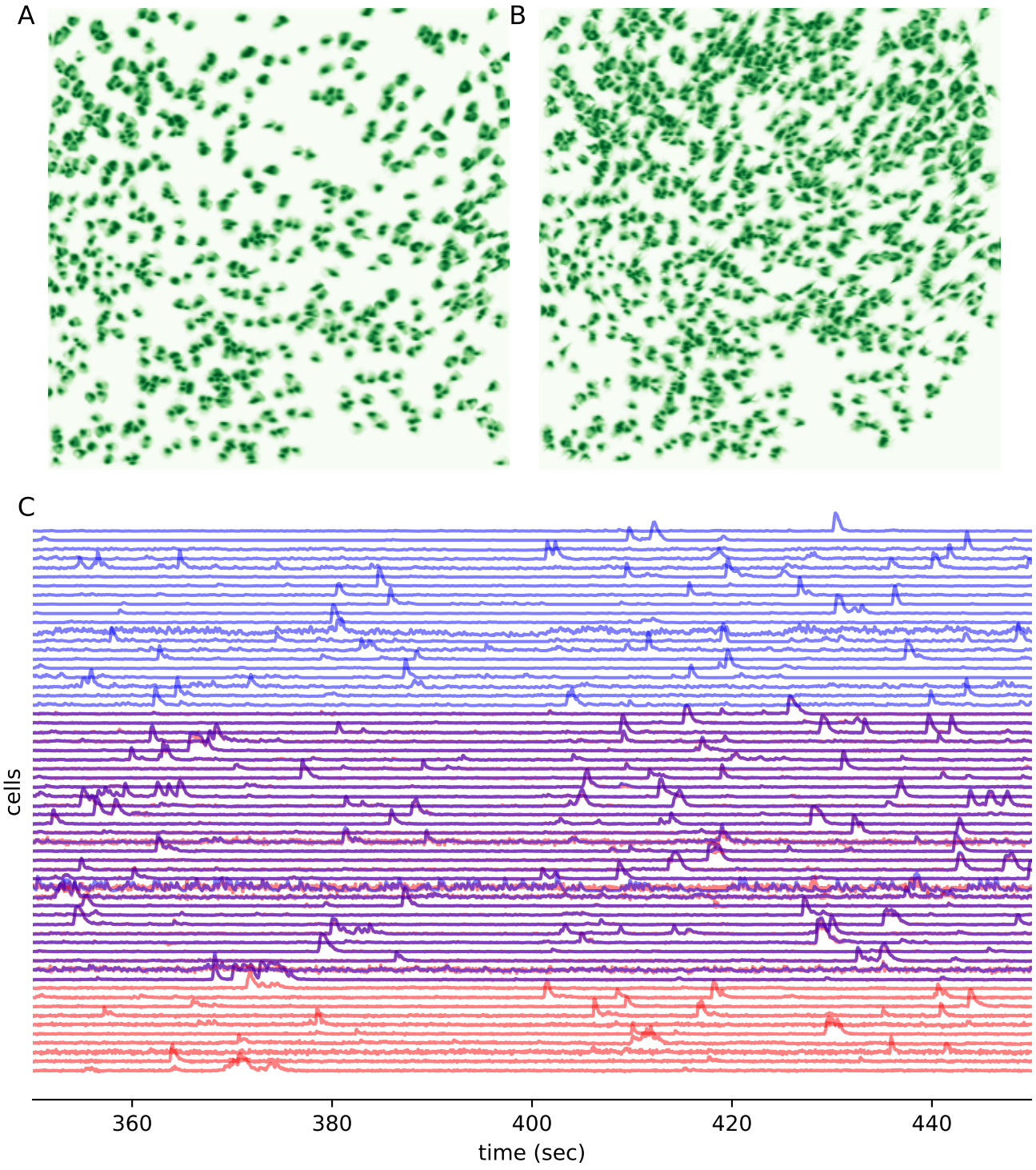

Figure 11. Comparison of CNMF-E and HOTARU on data 2: (A) footprints' contours by CNMF-E; (B) footprints by HOTARU; (C) part of estimated cell-by-cell calcium variability by CNMF-E (red) and HOTARU (blue). Overlapping areas indicate that the footprints of the candidate cells are similar.

Figure 11–video 1. Video comparing CNMF-E and HOTARU results for data 2.

Application of HOTARU to the data from hippocampus obtained by two-photon microscopy

We then analyzed the imaging data obtained by two-photon microscopy (experimental details and algorithm parameter settings for this data set are given in the Methods and Materials section). The data consists of 8989 frames of 512 × 512-pixel images of 532 × 532 μm taken at 15 Hz for 10 min. As with the data obtained by head-mount camera, we applied both HOTARU and CNMF-E to this data set to compare their performance.

In this data-set, a large number of small footprints were detected in the initial placement (***Figure 10***ABC). The small size of the footprints makes it difficult to distinguish neurons from false positives (***Figure 10***B). Based on the distribution of footprint radii in ***Figure 10***B with the first set of parameters, $r_{\max}$ was re-set to 9 and the neuron overlap criteria θ was set 1.6. After 10 iterations, the number of cells stabilized (***Figure 10***D), and a clean solidity distribution was obtained in the radius range of 2.2–6 (***Figure 10***E). The radius distribution of the detected cell candidates includes two groups, one centered at 2.7 pixels and the other at 4.5 pixels (***Figure Supplement 1***). They are thought to correspond to portions of dendrites and cell bodies, respectively.

Finally, 592 cells were detected by CNMF-E and 948 cells were detected by HOTARU, including some dendrites. Of the candidate cells, 506 were presumed identical because they had similar footprints and activity (***Figure 11***). As with previous data, the shape or activity pattern of the neurons detected with HOTARU only were undistinguishable from those detected by both methods (***Figure 11***).

Application of HOTARU to the large scale data from hippocampal CA3 obtained by miniature microscopy

Next, we analyze a large dataset of hippocampal data (***Figure 12***) (Full experimental details and algorithm parameter settings for this dataset are given in the Methods and Materials section). The video consists of 74000 frames of 607×344 pixel images taken at 20 Hz for 61 min. We applied both HOTARU and CNMF-E to the dataset and compared their performance with the previous data.

In the initial arrangement, the high-intensity cell candidates are concentrated in 4–8 pixels and appear to be easily distinguishable from false positives (***Figure 13***AB). We set $r_{\max}$ to 32 and θ to 1.6 to include the entire radius distribution. The cells recorded by miniature microscopy often appear blurred and larger due to peripheral distortion and depth of focus. Therefore, the value of $r_{\max}$ was set larger than the previous two sets of data. With the current data, many iterations were required for the cell counts to stabilize after the improvement iterations (***Figure 13***C). The final radius and firmness distributions obtained are sufficiently solidly distributed, as are the other data (***Figure 13***D).

Finally, 674 and 701 cell candidates were detected in CNMF-E and HOTRU, respectively (***Figure 14***). As in the previous two cases, 592 cell candidates were shown to be identical by CNMF-E and HOTARU. The number of cells detected is about the same, but the details of the candidate cells detected differ significantly (***Figure 14***AB). Many small cells are detected in the upper left region of the image when CNMF-E is used, whereas large cell candidates that may correspond to the distortion are detected when HOTARU is used. It is possible that the background removal in CNMF-E has produced a desirable result, but since no cell-like objects were detected in the upper left region of each frame from ***Figure 13***C, it is necessary to fully consider which result is closer to the truth. In addition, the footprints of the cells detected by CNMF-E are uniform in size, while the footprints detected by HOTARU are diverse in size and shape; CNMF-E may produce very different results when the cell size parameter is changed.

However, for this data, HOTARU shows some large size false positives corresponding to blood vessels and background. If $r_{\max}$ is reduced to eliminate false positives, even plausible cell candidates will not be detected. Therefore, it is desirable to increase $r_{\max}$ and eliminate obvious false positives. It is easier to identify large false positives that correspond to the background than small

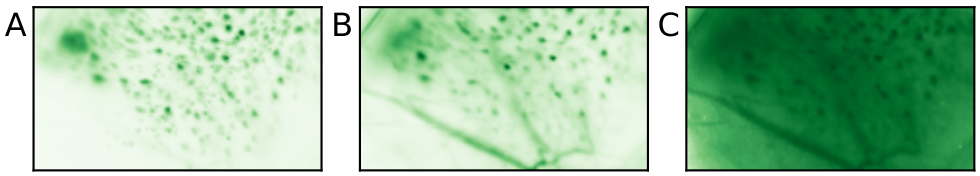

Figure 12. Data 3, large hippocampal CA3 data by miniature microscopy: the maximum (A), standard deviation (B), and local correlation (C).

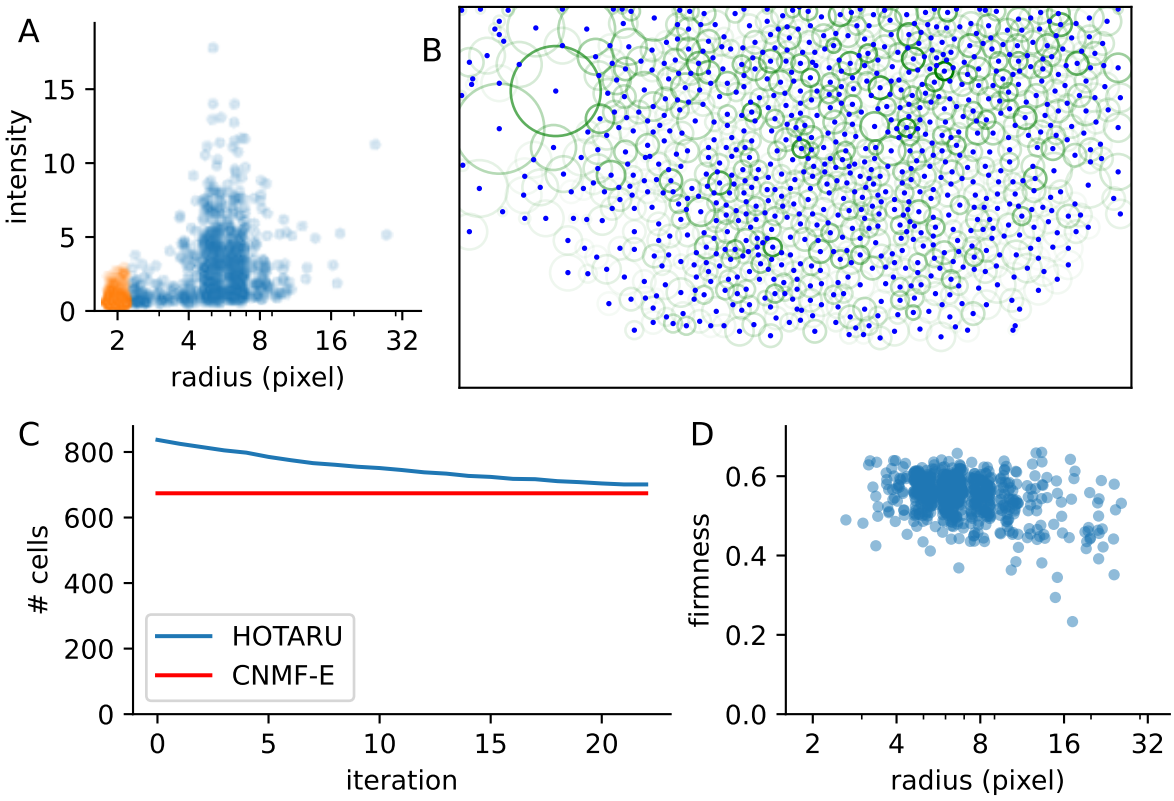

Figure 13. Summary of data 3: (A) distribution after cells are narrowed down using the greedy method with the cell-overlap criteria $\theta = 1.6$; (B) Overall view of the obtained peaks. (C) number of estimated cells after repeated improvement treatments. (D) distribution of cell radius and firmness at final step.

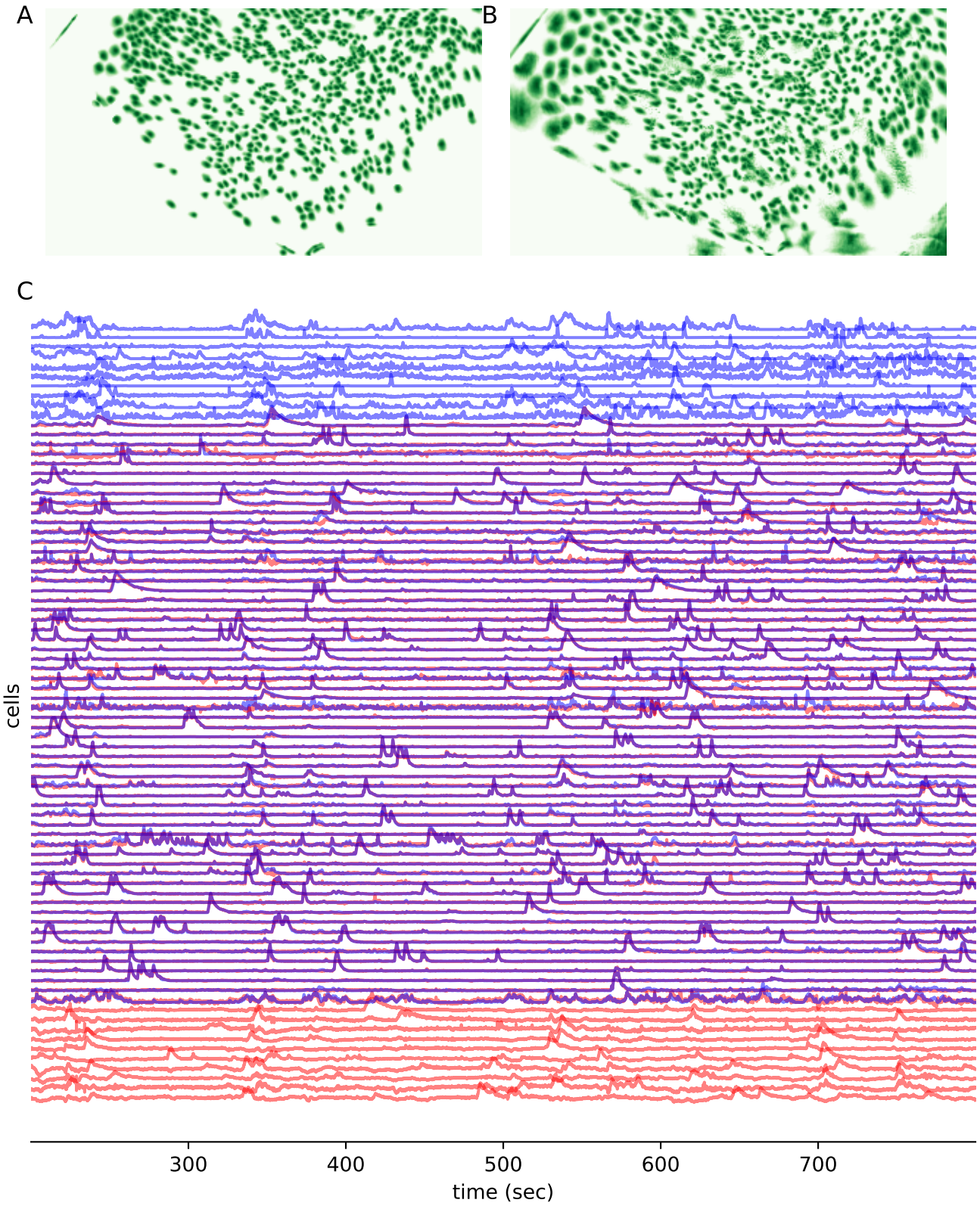

Figure 14. Comparison of CNMF-E and HOTARU on data 3: (A) footprints' contours by CNMF-E; (B) footprints by HOTARU; (C) part of estimated cell-by-cell calcium variability by CNMF-E (red) and HOTARU (blue). Overlapping areas indicate that the footprints of the candidate cells are similar.

Figure 14–video 1. Video comparing CNMF-E and HOTARU results for data 3.

false positives or excessive segmentation. HOTARU can also be used to mask out cell-free areas and areas on blood vessels from the calculation.

Although the ground truth is unknown, a large number of identical cells were detected by HOTARU and CNMF-E, and each is considered to be sufficiently reliable. Validation of multiple analyses by comparing them with each other continues to be an important issue.

Discussion

In neuroscience, calcium imaging is an important method for recording large neuronal populations, and CNMF for two-photon microscopy and CNMF-E for endoscopic microscopy enable highly accurate detection of cell activity even in large data sets. However, even in CNMF-E, parameter settings and algorithms can be arbitrary, making it difficult to verify whether the detected cells are truly correct, and there is a possibility that many detection errors occur in order to avoid false positives. Therefore, this paper proposes a simple, almost fully automatic cell activity detection system that is less arbitrary and less sensitive to parameter changes.

HOTARU is a very simple algorithm with few parameters, yet it can detect more cells with low firing frequency and low signal-to-noise ratio than CNMF-E. In addition, the fluorescence intensity estimates show a higher correlation with the true value in artificial data than CNMF-E. Detected cell candidates can be clearly distinguished from false positives by the distribution of two parameters: cell radius and solidity. These improvements should allow for more stable interpretation in downstream analyses. Conversely, in comparison to CNMF-E, there were cells detected only with CNMF-E and cells detected only with HOTARU. Objective evaluation of this result is difficult.

In this paper, we simply applied initialization and CNMF expansion, narrowed the cell removal criteria to cell radius shrinking, and did not perform any heuristic dividing, merging, or removal. This resulted in more iterations of footprint and spike improvements for larger data. The fact that very stable results were obtained despite a large number of iterations indicates the stability of the algorithm and increases confidence in the results. It should also be emphasized that the computation time per iteration is sufficiently small to allow such a verification.

HOTARU has six parameters: r_{max}, θ, λ_A, λ_U, b_T, and b_X. The appropriate value of r_{max} can be determined without arbitrariness, corresponding to the spread of the intensity distribution using a tentative value of r_{max}. Small value of θ will not only caouse a slight increase in the number of neurons that are eventually detected, but also increase the computation time. In practice, there is little difference in the number of cells for $\theta = 1.6$ and smaller values of θ. The regularization parameter for background b_T and b_X contribute to the stability of the optimization method. Based upon the effect of b_T and b_X on the obtained footprints and activities, the value of b_T and b_X can be set to 0.1. The larger value of λ_A detects fewer neurons but also results in fewer false positives. Setting $\lambda_A = 20$ almost eliminates false positives. Reducing λ_A may increase the number of neurons detected, but the change of number is small and does not result in a significant increase in false positives. The number of cells detected and the output footprint are not significantly affected by λ_U. Despite its impact on the accuracy of spike estimation, but it does not significantly affect the results. Furthermore, we can easily redo the spike estimation with a different λ_U after obtaining the final footprints. In conclusion, the analysis of artificial and experimental data shows that HOTARU can detect a large number of neurons with few false positives without arbitrary settings.

HOTARU is implemented in Python and TensorFlow and makes effective use of GPUs. The implementation is released as open source software (https://github.com/tk2lab/HOTARU), ver. 4.0, which corresponds to that used in this study. The implementation of HOTARU in this paper is sufficiently simple to evaluate the basic concept and is therefore easily extendable; there is room for further improvement by introducing local background processing and appropriate heuristics similar to those in CNMF-E.

Algorithm 1 Find initial peaks

Require: data $F[t,x]$, neuron size list $\{r_l\}_{l=1}^{N_r}$, minimal distance ratio θ
Ensure: peak information O

$L \leftarrow \{\}$
for $t \in \{1,2,\dots,T\}$ **do** ▷ find local maxima of z on x,l
  $z[t,l,x] \leftarrow F[t,x] * \mathrm{LoG}[r_l,x]$
  $Z[t,l,x] \leftarrow \mathrm{MaxmumFilter}_{x,l} z[t,l,x]$
  append $\{(t,l,x) \mid Z[t,l,x] = z[t,l,x]\}$ to L
end for
$P \leftarrow \{\}$
for all pixels x_* **do** ▷ reduce peaks (optional): $|P| \le X$
  $Q \leftarrow \{(t,l,x) \mid x = x_*,\ (t,l,x) \in L\}$
  if Q is not empty **then**
    append the element with the largest $z[t,l,x]$ in Q to P
  end if
end for
sort P in descending order by $z[t,l,x]$
$O \leftarrow \{\}$
while P is not empty **do** ▷ greedy selection
  $\hat{t},\hat{l},\hat{x} \leftarrow$ pop first element of P
  $\hat{r} \leftarrow r_{\hat{l}}$
  if $\forall (t,r,x) \in O, |\hat{x}-x| > \theta\hat{r}$ **then** ▷ check overlap
    append $(\hat{t},\hat{r},\hat{x})$ to O
    $P \leftarrow \{(t,l,x) \mid |\hat{x}-x| > \theta\hat{r},\ (t,l,x) \in P\}$
  end if
end while

Algorithm 2 Make intial footprints

Require: peak information $\hat{t},\hat{r},\hat{x}$
Ensure: radius $\hat{r}$, intensity $z_{\mathtt{max}}$, footprint $a[x]$

$z[x] \leftarrow F[\hat{t},x] * \mathrm{LoG}[\hat{r},x]$
$M \leftarrow$ connected region mask around peak $\hat{x}$ in image z
$z_{\mathrm{min}} \leftarrow \min_{x\in M} z[x]$
$z_{\mathrm{max}} \leftarrow \max_{x\in M} z[x]$ ▷ intensity
$a[x] \leftarrow (z[x]-z_{\mathrm{min}})/(z_{\mathrm{max}}-z_{\mathrm{min}})$ if $x \in M$ else 0 ▷ footprint

Algorithm 3 Clean up footprints

Require: dirty footprint $a[x]$, neuron size list $\{r_l\}_{l=1}^{N_r}$
Ensure: radius r, firmness g, clean footprint $a[x]$

$z[l,x] \leftarrow a[x] * \mathrm{LoG}[r_l,x]$
$\hat{l},\hat{x} \leftarrow \arg\max_{l,x} z[l,x]$
$M \leftarrow$ connected region around peak $\hat{x}$ in image $z[\hat{l}]$
$a_{\mathrm{min}} = \min_{x\in M} a[x]$
$a_{\mathrm{max}} = \max_{x\in M} a[x]$
$g \leftarrow z[\hat{l},\hat{x}]/(a_{\mathtt{max}} - a_{\mathtt{min}})$ ▷ firmness
$a[x] \leftarrow (a[x]-a_{\mathrm{min}})/(a_{\mathrm{max}}-a_{\mathrm{min}})$ if $x \in M$ else 0 ▷ clean footprint

Methods and Materials

Finding cell candidates from video data using LoG filters

To determine initial cell candidates, we first prepared a list of cell size parameters $\{r_l\}_{l=1}^{L}$. The LoG filter corresponding to all r_l was applied to the image $f_t[x]$ at time t to create L kinds of filtered images. This collection of images is considered to be a 3D structure of $L \times W \times H$ and all the maxima and their positions in the 3D space are collected.

Since the collected peak information is huge, it is efficiently condensed. First, the collected peaks are grouped by position x in the image, and only those with the largest maxima are retained. Concurrently, the cell size r of the adopted peak is also recorded. Next, the remaining peaks are sorted and adopted in descending order by the maximum value, i.e., signal intensity. However, if the distance between the centers is within θ times the radius of the peaks with stronger signals than those already selected, the peaks are not selected. The detailed procedure is shown in Algorithm 1.

Next, the cell shape is specifically determined for the obtained peak position and filter size. Once again, a LoG filter corresponding to the filter size is applied, and the watershed cut algorithm is applied to the center of the peak position to extract the cellular region. The filter image is cropped in the extracted region and normalized to $[0, 1]$ as the initial footprint. These details are shown in Algorithm 2.

The iterative footprint is also improved using a LoG filter and the watershed cut algorithm. Again, multiple LoG filters are applied to each footprint. The maximum peak is then found, and the watershed cut algorithm is applied to the filtered image around the peak to extract the cellular region. The cellular region is extracted from the original footprint and normalized to $[0, 1]$ to obtain the improved footprint. These details are shown in Algorithm 3.

Statistical Model

First, the time and space averages of video $F[t, x]$ should be set to zero:

$$F[t,x] \leftarrow F[t,x] - (M_t F)[x] - (M_x F)[t] + M_{tx} F, \tag{16}$$

and be normalized by the total standard deviation $F[t, x] \leftarrow F[t, x]/\mathrm{SD}[F]$.

To update the time variable $\hat{U} \in \mathbb{R}^{K\times T}$, the matrices $P \in \mathbb{R}^{K\times T}$, $Q \in \mathbb{R}^{K\times K}$, and $R \in \mathbb{R}^{K\times K}$ are computed a priori using the given $A_* \in \mathbb{R}^{K\times X}$ an its spatial mean $\overline{A}_* \in \mathbb{R}^{K\times 1}$:

$$P = F A_*^\top, \quad Q = A_* A_*^\top, \quad R = \overline{A}_* \overline{A}_*^\top. \tag{17}$$

To be evaluated, ***Equation 11*** can be computed using P, Q, and R as follows:

$$g[\hat{U}] = \left\| F - A_* \Gamma \hat{U} + B_\zeta A_* \hat{U} \Gamma \right\|_2^2 = |P \otimes \hat{V}|_1 + |Q \otimes C|_1 + |R \otimes O|_1 \tag{18}$$

where

$$\hat{V} = \hat{U}\Gamma, \quad C = \hat{V}^\top \hat{V}, \quad \overline{V} = M_t[\hat{V}], \quad O = \overline{V}^\top \overline{V}. \tag{19}$$

With the above preparations, we can find the $\hat{U}$ that minimizes L under given A_* by setting the initial value of $\hat{U}$ and repeatedly applying the proximal gradient method. The results in this paper use an initial value of 1 and apply Nesterev acceleration during iterations. An epoch is defined as 100 iterations and the iteration is terminated if the change in the value of L is less than tol before and after the epoch is performed.

A similar procedure is used for updating A. For given $V = U\Gamma$, we can define the evaluation function g as:

$$g[\hat{A}] = \left\| F - \hat{A}^\top V_* + B_\zeta \hat{A}^\top V_* \right\|_2^2 = |P \otimes \hat{A}|_1 + |Q \otimes C|_1 + |R \otimes O|_1, \tag{20}$$

$$P = F V_*^\top, \quad Q = V_* V_*^\top, \quad R = \overline{V}_* \overline{V}_*^\top, \tag{21}$$

$$C = \hat{A}^\top \hat{A}, \quad \overline{A} = M_x[\hat{A}], \quad O = \overline{A}^\top \overline{A}. \tag{22}$$

Table 2. Computational runnning time (sec) for processing the 4 datasets.

Dataset	Simulation Data	Data1	Data2	Data3
Size ($W \times H \times T$)	$200 \times 200 \times 1000$	$720 \times 540 \times 1200$	$512 \times 512 \times 8989$	$606 \times 344 \times 74000$
CNMF-E	35	289	870	11869
HOTARU	56	2371	5491	16540
(iteration)	(3)	(13)	(10)	(22)

Overall framework

The algorithm of the proposed method HOTARU is simple: (INIT) Algorithms 1 and 2 are used to obtain the initial footprints A; (T-P) temporal optimization is applied to obtain the spike sequence U; (S-P) spatial optimization is applied to obtain footprints A from the spike sequence U; (CLEAN) footprint A was modified by Algorithm 3. T-P, S-P, and CLEAN are repeated until the cell count stabilizes.

For all the analyses in the paper, there were 13 different parameters r, with r_{min} as the minimum value and r_{max} as the maximum value, for $\log r$ to be equally spaced. For all data, the r_{min} is fixed to 2. Conversely, to fit the actural cell size that can be estimated from the distribution of radius and firmness, the r_{max} were set to 20, 20, 9, and 32 respectively for the artificial and three real data (data1, data2, data3). The overlap level parameter θ was set to 1.2 for artificial data to evaluate the iteration effect and to 1.6 for real data. The response function to the spike was assumed to be $\gamma[\tau] = \exp(-t/\tau_{\text{rise}}) - \exp(-t/\tau_{\text{fall}})$ and $\tau_{\text{rise}} = 0.08$ sec, $\tau_{\text{fall}} = 0.16$ sec. In principle, the parameters of the normalization term were $\lambda_A = 20$, $\lambda_U = 30$, and $b_T = b_X = 0.1$, except for the evaluation in ***Figure 5***. The learning rate η of the proximal gradient method was set to 0.01 and the tolerance of the termination condition tol was set to 0.001.

Complexity analysis

Although calcium imaging is now being recorded on a larger scale, it is safe to assume that a certain upper limit is set for spatial resolution and that the computational complexity in the time direction T is most important. Considering the computational complexity of a time-step update, the matrix pre-computation is dominated by $O(KTX)$ due to the matrix product of F and A. In applying the proximal gradient, the computation of A dominates, $O(K^2T)$. For the space step, the pre-computation and the computational complexity per iteration respectively are $O(KTX)$ and $O(K^2X)$. Since the number of cells K is proportional to X but independent of T, the computational complexity for T is $O(T)$. For each frame, the process of finding the initial footprint can also be performed independently or in parallel, and by using per-pixel shrinkage, subsequent operations are independent of T.

Running time

Evaluate the computation time for the four datasets used in this study (***Table 2***). The details for the largest Data3 are: 269 sec for data normalization, 517 sec for peak discovery, 156 sec for peak refinement and contour cropping, 94 sec for first spike estimation, 629 sec for first footprint estimation, and 22 sec for cropping cell regions from footprints. Overall, CNMF-E is faster for small-scale data. HOTARU was run on an Intel Xeon W-2123 @ 3.60GHz with 128GB memory and Ubuntu 20.04.2 LTS on NVIDIA TITAN RTX. CNMF-E was run on a Windows OS on an Intel Xeon E5-2640 v4 @ 2.40 GHz with 128 GB memory. HOTARU was run until the cell counts were completely matched before and after the improvement iteration. It should be emphasized that changes in the later iterations were slight.

Simulated Data

Artificial data was created with 1,000 frames of 200 × 200 pixels (***Figure 1***). The cell radius r_i obeyes a uniform distribution in the range 4 to 8-pixel range, and the cell footprint was assumed to follow a two-dimensional Gaussian function of radius r_i. The cell centers were randomly placed under the condition that the distance between the centers of two adjacent cells was at least 1.8 times the radius r_i. Consequently, 181 cells ware located. The cell firing rates were set in the 0.2–2.0 Hz range, causing individual cells to fire from 6-123 times. The SN ratio of the cells' activity intensity obeyed uniform random distribution in the 0.5–2.5 range.

The background activities was calculated as follows:

$$b[t,x] = -5\frac{\|x - \texttt{center}\|^2}{R^2} + 0.5\cos\frac{4\pi t}{T} + \varepsilon[t,x], \quad \varepsilon[t,x] \sim \mathcal{N}(0,1^2) \tag{23}$$

Subject Details

All procedures involving the use of animals complied with the guidelines of the National Institutes of Health, and they were approved by the Animal Care and Use Committee of the University of Toyama and the Institutional Committee for the Care and RIKEN Animal Experiments Committee and Genetic Recombinant Experiment Safety Committee. Thy1::G-CaMP7-T2A-DsRed2 mice that express the fluorescent calcium indicator protein G-CaMP7 and the red fluorescent protein DsRed2 under the control of the neuron-specific Thy1 promoter were used. Details on the generation and characterization of the transgenic mice are described in ***Sato et al.*** (***2020***).

Hippocampal data by miniature microscope

The mice were kept on a 12-h/12-h light-dark cycle (lights on 7:00 am) at 24 ± 3° C and 55 ± 5% humidity, provided with food and water ad libitum, and housed with littermates until 1—5 days before surgery. We performed hippocampal surgery for gradient refractive index (GRIN) lens setting as previously described (***Barretto et al., 2011***; ***Ghosh et al., 2011***; ***Ziv et al., 2013***; ***Ghandour et al., 2019***). All surgery was conducted on approximately 12-week-old male Thy1::G-CaMP7-p2A-DsRed mice on a C57BL/6J background. The mice were anesthetized with a pentobarbital solution (80 mg/kg of body weight; intraperitoneal injection), and the fully anesthetized mice were placed in a stereotactic apparatus (Narishige, Japan). To set the cannula lens sleeve (outer diameter, 1.8 mm; length, 3.6 mm; Inscopix, CA), craniotomy was performed with a diameter of 2.0 mm. The cylindrical column of the neocortex and corpus callosum above the alveus covering the dorsal hippocampus was aspirated using a 27-gauge blunt drawing up needle with saline. The cannula lens sleeve was gently placed on the alveus and fixed to the edge of the burr hole with bone wax, which was melted using a low-temperature cautery. The cannula lens sleeve targeted the right hemisphere (AP 2.0 mm, ML 1.5 mm at center). After setting the anchor screws onto the skull, we covered the skull with dental cement, which fixed the cannula lens sleeve to the skull and anchor screws.

Approximately 3–4 weeks after surgery, mice were anesthetized with isoflurane (1.5–2%), and a GRIN lens (outer diameter, 1.0 mm; length, 4.0 mm; Inscopix) was inserted into the cannula lens sleeve and fixed with ultraviolet-curing adhesive (Norland, NOA 81). The integrated miniature microscope (nVista HD, Inscopix) (***Ghosh et al., 2011***) with a microscope baseplate (Inscopix) was placed above the GRIN lens, at which G-CaMP7 fluorescence was observed. The microscope baseplate was fixed with the head of the anchor screw using dental cement. This shaded the GRIN lens, after which the integrated microscope was detached from the baseplate. The GRIN lens was covered by attaching the microscope baseplate cover to the base plate until calcium imaging was performed.

Calcium imaging was performed during the light cycle, and calcium events were captured at 20 Hz using nVista acquisition software (Inscopix). The integrated microscope was re-attached to the baseplate, and then the mice were introduced into a novel context consisting of a cylindrical chamber (diameter × height: 180 × 230 mm^2) with a white acrylic floor and walls covered with black

tape. The captured video was processed as previously described (***Kitamura et al., 2015***; ***Ghandour et al., 2019***). Motion correction by Mosaic software (Inscopix) was applied to th dF/F signal video.

Two-photon calcium imaging

The two-photon microscopy images used in ***Figure 9***, ***Figure 10***, and ***Figure 11*** were acquired using a Nikon A1MP (Nikon, Tokyo, Japan) equipped with a 16x, NA 0.8 water immersion objective lens as described previously (***Sato et al., 2020***). Briefly, an optical window was surgically implanted in cortex overlaying dorsal hippocampus. G-CaMP7 and DsRed2 were excited at 910 nm using a Ti-sapphire laser (MaiTai DeepSee eHP, Spectra-Physics, Milpitas, CA). G-CaMP7 fluorescence was separated using a dichroic mirror at 560 nm and collected using an external GaAsP photomultiplier tube (10770PB-40, Hamamatsu Photonics, Japan). The microscope focused at approximately 150 µm from the hippocampal surface to image CA1 pyramidal cells labeled with G-CaMP7. A resonant-galvo scanner attached to the microscope acquired 512×512 pixel images at 15 frames per second. An imaging session lasted 10 minutes. The field of view was 532×532 µm.

Acknowledgments

This work was supported by JSPS KAKENHI (grant number: JP18H05213), the Core Research for Evolutional Science and Technology (CREST) program (JPMJCR13W1) of the Japan Science and Technology Agency (JST), a Grant-in-Aid for Scientific Research on Innovative Areas “Memory dynamism” (JP25115002) from MEXT, the Takeda Science Foundation support to K.I. T.F. is supported by the Brain/Minds project from the MEXT, KAKENHI (JP15H04265) from the JSPS, a Grant-in-Aid for Scientific Research on Innovative Areas “Memory dynamis” (JP16H01289) and “Artificial Intelligence and Brain Science” (JP17H06036) from the MEXT. Y.H. is supported by RIKEN, the Human Frontier Science Programme (RGP0022/2013), the Core Researh for Evolutional Science and Technology (CREST) program (JPMJCR20E4) of the Japan Science and Technology Agency (JST), a Grant-in-Aid for Scientific Research on Innovative Areas “Memory dynamism” (JP16H01292), “Synapse Pathology” (JP22110006), “Multiscale Brain” (JP18H05434) from the MEXT. T.T. is supported by KAKENHI (JP26870577, JP19K12104) from the JSPS.

Competing interests

Y.H. received research fund from Takeda Pharmaceuticals, Fujitsu Laboratories, and Dwango.

References

Ahrens MB, Orger MB, Robson DN, Li JM, Keller PJ. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nature methods. 2013; 10(5):413–420.

Barbera G, Liang B, Zhang L, Gerfen CR, Culurciello E, Chen R, Li Y, Lin DT. Spatially compact neural clusters in the dorsal striatum encode locomotion relevant information. Neuron. 2016; 92(1):202–213.

Barretto RP, Ko TH, Jung JC, Wang TJ, Capps G, Waters AC, Ziv Y, Attardo A, Recht L, Schnitzer MJ. Time-lapse imaging of disease progression in deep brain areas using fluorescence microendoscopy. Nature medicine. 2011; 17(2):223–228.

Chen TW, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, Schreiter ER, Kerr RA, Orger MB, Jayaraman V, et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature. 2013; 499(7458):295–300.

Cotton RJ, Froudarakis E, Storer P, Saggau P, Tolias AS. Three-dimensional mapping of microcircuit correlation structure. Frontiers in neural circuits. 2013; 7:151.

Cousty J, Bertrand G, Najman L, Couprie M. Watershed cuts: Minimum spanning forests and the drop of water principle. IEEE transactions on pattern analysis and machine intelligence. 2008; 31(8):1362–1374.

Diego Andilla F, Hamprecht FA. Learning multi-level sparse representations. Advances in Neural Information Processing Systems. 2013; 26.

Flusberg BA, Nimmerjahn A, Cocker ED, Mukamel EA, Barretto RP, Ko TH, Burns LD, Jung JC, Schnitzer MJ. High-speed, miniaturized fluorescence microscopy in freely moving mice. Nature methods. 2008; 5(11):935–938.

Ghandour K, Ohkawa N, Fung CCA, Asai H, Saitoh Y, Takekawa T, Okubo-Suzuki R, Soya S, Nishizono H, Matsuo M, et al. Orchestrated ensemble activities constitute a hippocampal memory engram. Nature communications. 2019; 10(1):2637.

Ghosh KK, Burns LD, Cocker ED, Nimmerjahn A, Ziv Y, El Gamal A, Schnitzer MJ. Miniaturized integration of a fluorescence microscope. Nature methods. 2011; 8(10):871–878.

Kitamura T, Sun C, Martin J, Kitch LJ, Schnitzer MJ, Tonegawa S. Entorhinal cortical ocean cells encode specific contexts and drive context-specific fear memory. Neuron. 2015; 87(6):1317–1331.

Klaus A, Martins GJ, Paixao VB, Zhou P, Paninski L, Costa RM. The spatiotemporal organization of the striatum encodes action space. Neuron. 2017; 95(5):1171–1180.

Lindeberg T. Scale selection properties of generalized scale-space interest point detectors. Journal of Mathematical Imaging and vision. 2013; 46(2):177–210.

Maruyama R, Maeda K, Moroda H, Kato I, Inoue M, Miyakawa H, Aonishi T. Detecting cells using non-negative matrix factorization on calcium imaging data. Neural Networks. 2014; 55:11–19.

Mukamel EA, Nimmerjahn A, Schnitzer MJ. Automated analysis of cellular signals from large-scale calcium imaging data. Neuron. 2009; 63(6):747–760.

Nesterov Y. A method for unconstrained convex minimization problem with the rate of convergence O (1/kˆ 2). In: *Doklady an ussr*, vol. 269; 1983. p. 543–547.

Nesterov Y. On an approach to the construction of optimal methods of minimization of smooth convex functions. Ekonomika i Mateaticheskie Metody. 1988; 24(3):509–517.

Pinto L, Dan Y. Cell-type-specific activity in prefrontal cortex during goal-directed behavior. Neuron. 2015; 87(2):437–450.

Pnevmatikakis EA, Merel J, Pakman A, Paninski L. Bayesian spike inference from calcium imaging data. In: *2013 Asilomar Conference on Signals, Systems and Computers* IEEE; 2013. p. 349–353.

Pnevmatikakis EA, Soudry D, Gao Y, Machado TA, Merel J, Pfau D, Reardon T, Mu Y, Lacefield C, Yang W, et al. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. Neuron. 2016; 89(2):285–299.

Prevedel R, Yoon YG, Hoffmann M, Pak N, Wetzstein G, Kato S, Schrödel T, Raskar R, Zimmer M, Boyden ES, et al. Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. Nature methods. 2014; 11(7):727–730.

Resendez SL, Jennings JH, Ung RL, Namboodiri VMK, Zhou ZC, Otis JM, Nomura H, McHenry JA, Kosyk O, Stuber GD. Visualization of cortical, subcortical and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. Nature protocols. 2016; 11(3):566–597.

Sato M, Mizuta K, Islam T, Kawano M, Sekine Y, Takekawa T, Gomez-Dominguez D, Schmidt A, Wolf F, Kim K, et al. Distinct mechanisms of over-representation of landmarks and rewards in the hippocampus. Cell reports. 2020; 32(1):107864.

Smith SL, Häusser M. Parallel processing of visual space by neighboring neurons in mouse visual cortex. Nature neuroscience. 2010; 13(9):1144–1149.

Takekawa T, Asai H, Ohkawa N, Nomoto M, Okubo-Suzuki R, Ghandour K, Sato M, Hayashi Y, Inokuchi K, Fukai T. Automatic sorting system for large calcium imaging data. BioRxiv. 2017; p. 215145.

Zhou P, Resendez SL, Rodriguez-Romaguera J, Jimenez JC, Neufeld SQ, Giovannucci A, Friedrich J, Pnevmatikakis EA, Stuber GD, Hen R, et al. Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. Elife. 2018; 7:e28728.

Ziv Y, Burns LD, Cocker ED, Hamel EO, Ghosh KK, Kitch LJ, El Gamal A, Schnitzer MJ. Long-term dynamics of CA1 hippocampal place codes. Nature neuroscience. 2013; 16(3):264–266.

Ziv Y, Ghosh KK. Miniature microscopes for large-scale imaging of neuronal activity in freely behaving rodents. Current opinion in neurobiology. 2015; 32:141–147.

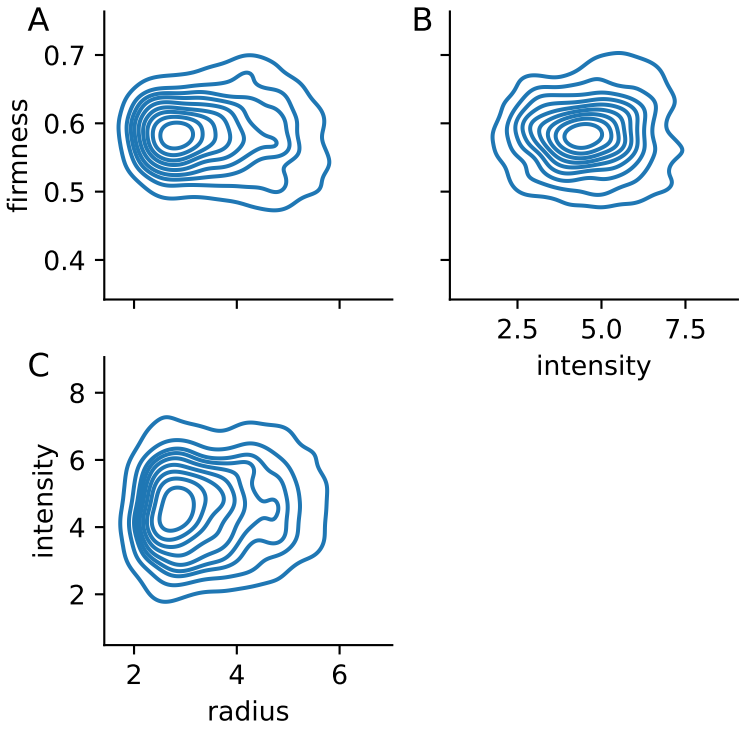

Figure 10–Figure supplement 1. Detailed distribution of radius and firmness and activity intensity s_k.