UC San Diego
UC San Diego Previously Published Works

Title

PatchWarp: Corrections of non-uniform image distortions in two-photon calcium imaging data by patchwork affine transformations

Permalink

https://escholarship.org/uc/item/4jh6w2sx

Journal

Cell Reports Methods, 2(5)

ISSN

2667-2375

Authors

Hattori, Ryoma

Komiyama, Takaki

Publication Date

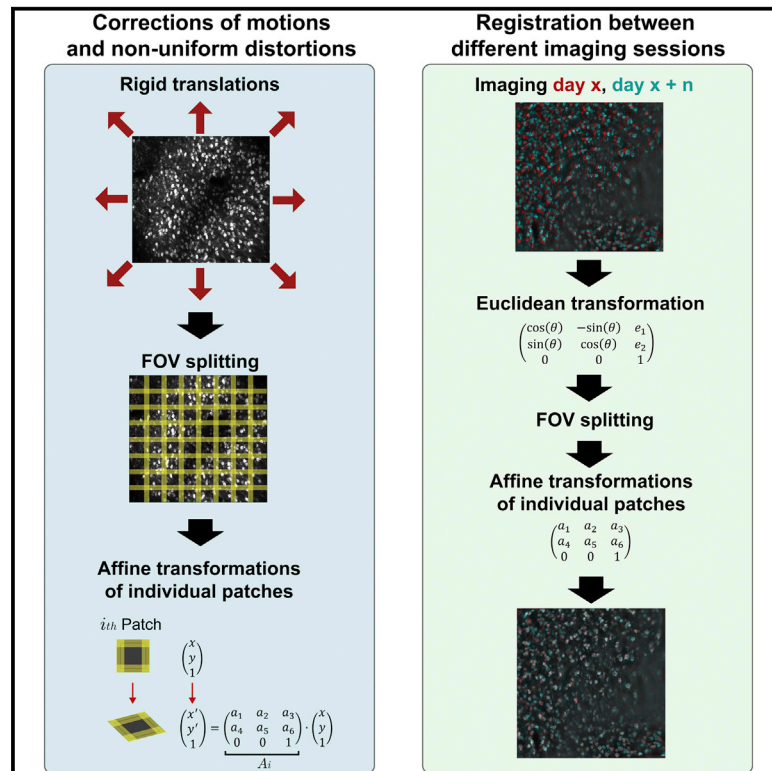2022-05-01

DOI

10.1016/j.crmeth.2022.100205

Copyright Information

Peer reviewed

Article

PatchWarp: Corrections of non-uniform image distortions in two-photon calcium imaging data by patchwork affine transformations

Graphical abstract

Authors

Ryoma Hattori, Takaki Komiyama

Correspondence

rhattori0204@gmail.com (R.H.),
tkomiyama@ucsd.edu (T.K.)

In brief

Hattori and Komiyama develop an image-registration algorithm for correcting complex image distortions in calcium imaging data. It outperforms existing non-rigid registration algorithms and is successful at registering images from different imaging sessions. The approach would be useful to correct image distortions in a wide range of disciplines.

Highlights

- PatchWarp registers images based on patchwork affine transformations

- PatchWarp corrects motion artifacts and complex distortions on calcium imaging data

- PatchWarp achieves superior performance in accuracy and efficiency

- PatchWarp also allows automated across-session registration for longitudinal analyses

CelPress

PatchWarp: Corrections of non-uniform image distortions in two-photon calcium imaging data by patchwork affine transformations

Ryoma Hattori[1,2,3,*] and Takaki Komiyama[1,2,*]

[1]Neurobiology Section, Center for Neural Circuits and Behavior, Department of Neurosciences, University of California San Diego, La Jolla, CA 90093, USA
[2]Halıcıoğlu Data Science Institute, University of California San Diego, La Jolla, CA 90093, USA
[3]Lead contact
*Correspondence: rhattori0204@gmail.com (R.H.), tkomiyama@ucsd.edu (T.K.)
https://doi.org/10.1016/j.crmeth.2022.100205

MOTIVATION Two-photon microscopy has been widely used to record the activity of populations of individual neurons at high spatial resolution in behaving animals. The ability to perform imaging for an extended period of time allows the investigation of activity changes associated with behavioral states and learning. However, imaging often accompanies shifts of the imaging field, including rapid (∼100 ms) translation and slow, spatially non-uniform distortion. To combat this issue and obtain a stable time series of the target structures, motion correction algorithms are commonly applied. However, typical motion correction algorithms for calcium imaging are limited to translation of images or image subfields and are unable to correct more complex non-uniform distortions.

SUMMARY

Complex distortions on calcium imaging often impair image registration accuracy. Here, we developed a registration algorithm, PatchWarp, to robustly correct slow image distortion for calcium imaging data. PatchWarp is a two-step algorithm with rigid and non-rigid image registrations. To correct non-uniform image distortions, it splits the imaging field and estimates the best affine transformation matrix for each of the subfields. The distortion-corrected subfields are stitched together like a patchwork to reconstruct the distortion-corrected imaging field. We show that PatchWarp robustly corrects image distortions of calcium imaging data collected from various cortical areas through glass window or gradient-index (GRIN) lens with a higher accuracy than existing non-rigid algorithms. Furthermore, it provides a fully automated method of registering images from different imaging sessions for longitudinal neural activity analyses. PatchWarp improves the quality of neural activity analyses and is useful as a general approach to correct image distortions in a wide range of disciplines.

INTRODUCTION

Two-photon microscopy has become a popular technique in neuroscience (Svoboda and Yasuda, 2006). It allows imaging of neuronal structures and activity inside the intact brain at high spatial resolution. The most common approach to measure neural activity with two-photon microscopy is calcium imaging. Calcium imaging is based on the principle that the intracellular concentration of calcium ion increases when neurons increase their firing rate. Therefore, calcium imaging allows an indirect measurement of neuronal spiking activity using fluorescent calcium indicators (Grienberger and Konnerth, 2012). Two-photon calcium imaging has several advantages over the classical techniques of recording neural activity using electrodes. For

example, calcium imaging allows simultaneous recording of hundreds of neurons within a small area. It also allows longitudinal tracking of the activity of those neurons over weeks and months. Furthermore, specific cell types can be identified by targeting the expression of genetically encoded calcium indicators (GECIs) with viral or transgenic approach.

However, two-photon calcium imaging suffers from motion artifacts, especially when it is applied to behaving animals. Such motion artifacts must be removed to accurately quantify neural activity. As such, a number of algorithms have been introduced to correct motion artifacts of two-photon calcium imaging data. Fast-motion (∼100 ms) artifacts mostly reflect spatially uniform translations as long as each frame is scanned quickly within tens of milliseconds. These translations can be efficiently

corrected by rigid alignment of individual image frames to a template image (Dubbs et al., 2016; Guizar-Sicairos et al., 2008; Mitani and Komiyama, 2018; Pachitariu et al., 2017; Pnevmatikakis and Giovannucci, 2017; Thévenaz et al., 1998). These rigid alignment methods have been extensively used for post hoc processing of two-photon calcium imaging data.

In recent years, it has become possible to perform calcium imaging for a long period of time, even for a continuous few hours without photobleaching due to improvement in calcium indicators (Chen et al., 2013; Dana et al., 2019; Inoue et al., 2019; Zhang et al., 2021). Improved stability and photosensitivity of calcium indicators allow stable calcium imaging with minimum laser power. Continuous, long calcium imaging experiments are often necessary to understand the neural mechanisms of cognition, state representation, and learning (Hattori and Komiyama, 2022; Hattori et al., 2017, 2019). However, rigid alignments of motion artifacts are often not sufficient for such long imaging sessions, because the image fields can slowly distort over time. The slow image distortions can be caused by both biological and mechanical issues. For example, the shape of brain tissue may slowly change due to behaviors, metabolic activity, dilation of blood vessels, and liquid-reward consumptions. Furthermore, mechanical instability of the animal stage or the microscope can also introduce slow non-uniform image distortions, particularly when the angle of the focus plane drifts over time. These image distortions limit the time of stable continuous imaging. Existing non-rigid motion-correction algorithms (Chen et al., 2019; Giovannucci et al., 2019; Pachitariu et al., 2017; Pnevmatikakis and Giovannucci, 2017) partially mitigate this problem by splitting an imaging field of view (FOV) and estimating the optimal translational shifts for the multiple subfields, but such subfieldwise rigid translations are suboptimal to correct non-uniform distortions.

To overcome this limitation of two-photon calcium imaging, we developed a novel algorithm for the processing of two-photon calcium imaging data. The algorithm, PatchWarp, is a two-step algorithm with rigid corrections and warp corrections. First, PatchWarp corrects uniform motion artifacts by performing rigid alignments of image frames to a template image by iteratively re-estimating the template image. Then, it corrects image distortions by performing affine transformations on the subfields of each image frame. We show that PatchWarp effectively corrects image distortions of two-photon calcium imaging data of both cell bodies and axons, either through a glass window or an implanted gradient-index (GRIN) lens. Furthermore, we show that PatchWarp can also register images from different imaging sessions for longitudinal across-session analyses of neural activity. We openly share the software at https://github.com/ryhattori/PatchWarp.

RESULTS

Non-uniform, slow image distortions during *in vivo* two-photon calcium imaging

We first present example two-photon calcium imaging sessions with non-uniform image distortions. We used a cell-body-imaging session (∼3.5 h of imaging; retrosplenial cortex; camk2-tTA:tetO-GCaMP6s transgenic mouse; Mayford et al., 1996; Wekselblatt

et al., 2016) and an axon-imaging session (∼20 min of imaging; primary motor cortex; AAV-hSyn-FLEX-axon-GCaMP6s; Broussard et al., 2018; in the basal forebrain of Chat-Cre transgenic mouse; Rossi et al., 2011) as the examples. We compared the spatial positions of neuronal structures between early and late frames of each imaging session after correcting motion artifacts by rigid motion corrections (Figure 1). These examples clearly indicate the presence of non-uniform image distortions during imaging. Only some subfields are registered between early and late frames with a rigid motion-correction algorithm. Corrections of these image distortions are of critical importance for the analysis of two-photon calcium imaging data because a region of interest (ROI) drawn on a neuronal structure (e.g., cell body, axon bouton, and spine) will be slowly displaced from the target structure in the presence of image distortions.

PatchWarp algorithm for non-rigid image registration with distortion correction

We developed a novel algorithm, PatchWarp, to robustly correct both rigid translation and non-rigid distortion on two-photon calcium imaging data. PatchWarp is a two-step algorithm that consists of rigid corrections and warp corrections (Figure 2A).

The rigid motion correction was done by identifying the translational shift that maximizes the correlation between a template image and each image frame. We estimated the optimal translational shifts with subpixel accuracy using a pyramid-based hill-climbing algorithm (Adelson et al., 1984; Mitani and Komiyama, 2018). We found that, in some cases, especially during long imaging sessions, the net slow image distortions over the course of an imaging session can be severe. In these cases of severe non-rigid distortions, a single template image is often not sufficient to accurately estimate translations of all image frames over time because of the large difference between the template image and the distorted frames. Therefore, PatchWarp iteratively updates the template image during the rigid motion correction (Figures 2B and 2C). The first template image is the average of the frames around the middle of a session. This template is used to motion correct the middle frames, and the template is updated by averaging the motion-corrected middle frames. Then, the motion correction proceeds bidirectionally from the middle toward both ends. Along the way, the template image is iteratively updated to reflect slow image distortions. This iterative template updating makes the rigid motion correction robust on imaging sessions with slow image distortions.

The rigid motion-correction step is followed by a warp correction. As the examples in Figure 1 show, the image distortions on two-photon calcium imaging data can be non-uniform across the field of view (FOV). This non-uniformity makes it impossible for a single geometric transformation function to fully register the images. Therefore, we implemented a "patchwork" approach where a FOV is split into subfields and the optimal transformations are separately estimated for each subfield (Figure 2D). The number of subfields can be adjusted for each imaging condition based on its distortion pattern. We transformed the coordinate of each patch using affine transformation. Affine transformation is a geometric transformation that preserves collinearity and ratios of distances. It can transform an image by translation, rotation, scaling, and shearing. Affine
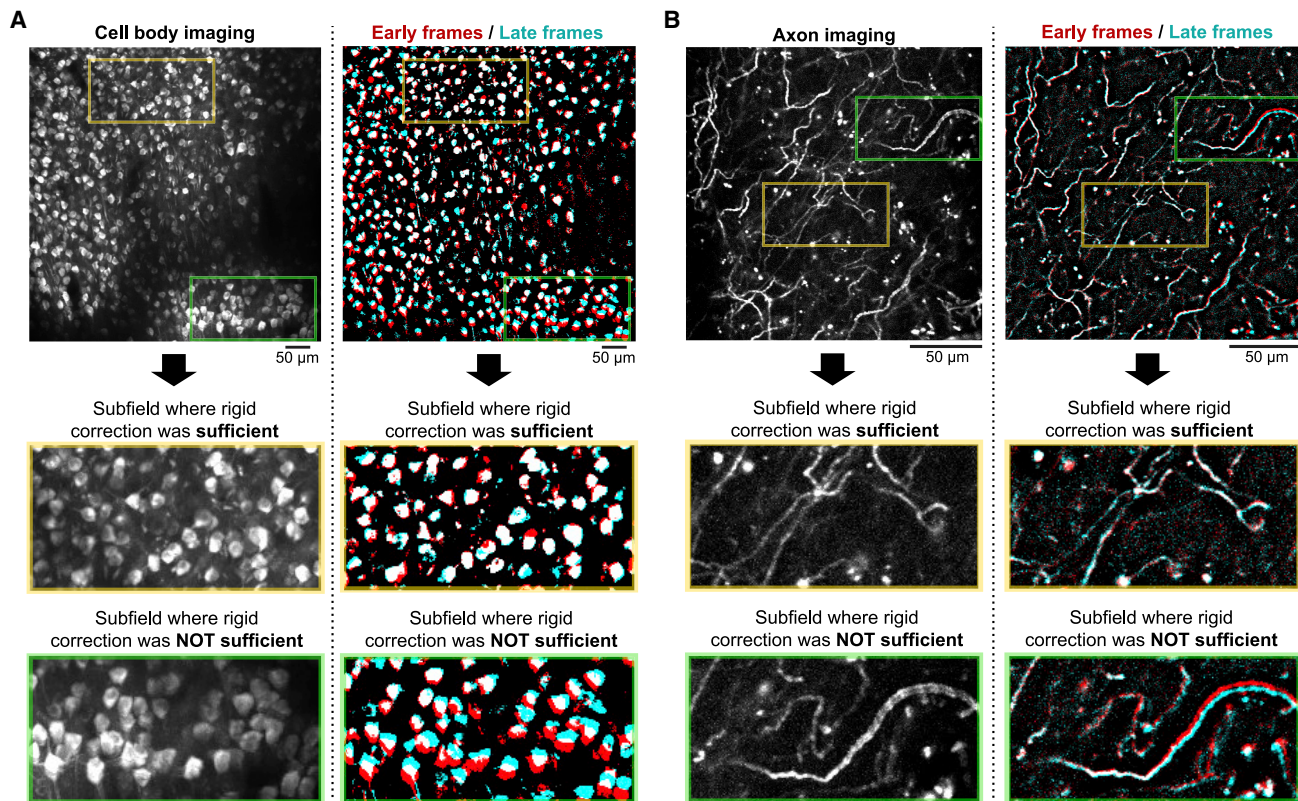
Figure 1. Non-uniform slow image distortions during *in vivo* two-photon calcium imaging

(A) Max-intensity projection image of an example RSC-imaging session. The max-intensity projection of all frames after correcting only rigid motion artifacts (left) and the max-intensity projections of the first (red) and the last (cyan) 14-min frames in the 3.5-h session after correcting only rigid motion artifacts (right) are shown. The second and third rows show the zoomed images that highlight the subfields where rigid corrections sufficiently (second row) or insufficiently (third row) registered early and late frames. The colors on the right panels are intentionally saturated with an arbitrary threshold to highlight the image overlaps.

(B) Max-intensity projection image of an example cholinergic axon-imaging session. The max-intensity projection of all frames after correcting only rigid motion artifacts (left) and the max-intensity projections of the first (red) and the last (cyan) 1.4-min frames in the 20-min session after correcting only rigid motion artifacts (right) are shown.

transformation of an image is done by a matrix multiplication between the original image coordinate and an affine transformation matrix. The original image can be transformed to match the transformed coordinate by subpixel bilinear interpolation such that the image spatially fits the new coordinate. By using the optimal affine transformation matrix for each patch, we can correct non-uniform image distortions that exist in the original full FOV. The optimal affine transformation matrix of each patch can be estimated by a gradient-based algorithm that aims at finding the optimal transformation matrix that maximizes the enhanced correlation coefficient (ECC) between a template image and an affine-transformed target image (Evangelidis and Psarakis, 2008; Psarakis and Evangelidis, 2005). We created the template image by averaging motion-corrected frames around the middle of a session (Figures 2E and 2F). Unlike the rigid motion-correction step, where the template image is iteratively updated, we use only a single template for all frames in a session for the second step of non-rigid warp correction to map all frames onto a single reference geometry. This is possible because of the flexible geometric transformation by the patchwork affine transformations. The source images to calculate

the transformation matrix were created by downsampling all frames by non-overlapping moving averaging of 500 frames. Therefore, the total number of affine transformations that we need to estimate is (no. of patches) × (no. of downsampled frames).

One remaining issue is that different neurons within each patch change calcium signals differently across time (i.e., the relative brightness across different cells is dynamic), which can affect the correlation between the template and the target images. To make the estimations of optimal transformations robust against the variability of calcium levels across neurons, we normalized the intensity of each pixel by dividing the pixel value by the mean intensity of the nearby pixels (Figures 2E and S1). This local intensity normalization allows the gradient-based algorithm to focus only on the cellular structures while ignoring calcium dynamics.

The PatchWarp pipeline corrected the non-uniform image distortions of the example sessions from Figure 1 (Figure S2).

Performance of PatchWarp algorithm
To examine the performance and general applicability of the PatchWarp algorithm, we used three different *in vivo* two-photon
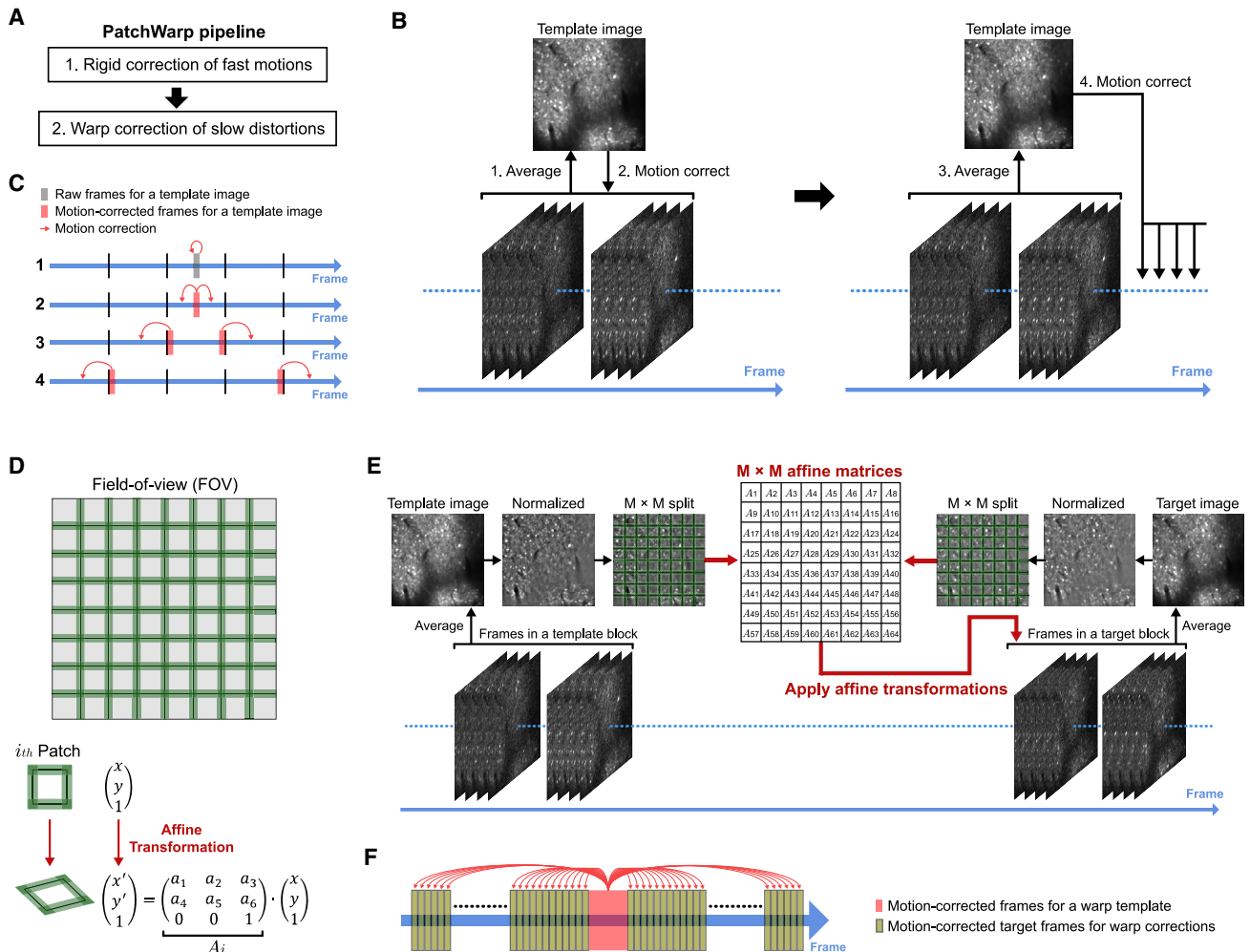
Figure 2. PatchWarp algorithm for correcting rigid motion artifacts and slow image distortions

(A) PatchWarp is a two-step algorithm. The first step corrects the rigid motion artifacts, and the second step corrects the slow image distortions.

(B) Schematic of the rigid motion correction. The first template image is created by averaging middle frames in a session, and these middle frames are registered to the mean image. The template image is updated by averaging the registered middle frames, and the updated template is used to register other frames.

(C) Timeline of the rigid motion correction. The rigid motion correction proceeds bidirectionally from the middle of a session. The template images are iteratively updated after specified frame number to reflect slow distortions on the template images.

(D) Warp correction is done by splitting a FOV into small overlapping patches and finding the optimal affine transformation matrix for each patch. Each patch is transformed by subpixel bilinear interpolation to fit the transformed coordinate.

(E) Schematic of the warp correction. Frames in the template and the target blocks are averaged, and the mean images are normalized by local intensity. The normalized images are split into patches, and an affine transformation that maximizes ECC between the template image and the transformed target image is estimated for each patch. The obtained transformations are applied to individual frames in the target block.

(F) Timeline of the warp correction. Warp correction uses a fixed-template image that is created by averaging the middle frames of a session. Optimal affine transformations are independently estimated for each target block. The target block size should be determined based on the distortion speed in the imaging data. See also Figures S1 and S2.

calcium imaging conditions that were recorded from awake-behaving mice. The first condition is the imaging of neuronal cell bodies from six different dorsal cortical areas (ALM, anterior-lateral motor area; pM2, posterior secondary motor cortex; PPC, posterior parietal cortex; RSC, retrosplenial cortex; S1, primary somatosensory cortex; V1, primary visual cortex; n = 1 session from each area; camk2-tTA:tetO-GCaMP6s transgenic mice) through glass windows. The second condition is the imaging of neuronal cell bodies from a ventral cortical area (OFC,

orbitofrontal cortex; n = 6; camk2-tTA:tetO-GCaMP6s transgenic mice) through GRIN lenses that were implanted in the brains. The third condition is the imaging of axons from a dorsal cortex (M1, axons of basal forebrain cholinergic neurons imaged in the primary motor cortex; n = 6; AAV-hSyn-FLEX-axon-GCaMP6s in Chat-Cre transgenic mice) through glass windows.

We used two metrics to quantify the registration performance, "mean max-intensity difference (mMD)" and "mean correlation with mean image (mCM)." To calculate mMD, the maximum
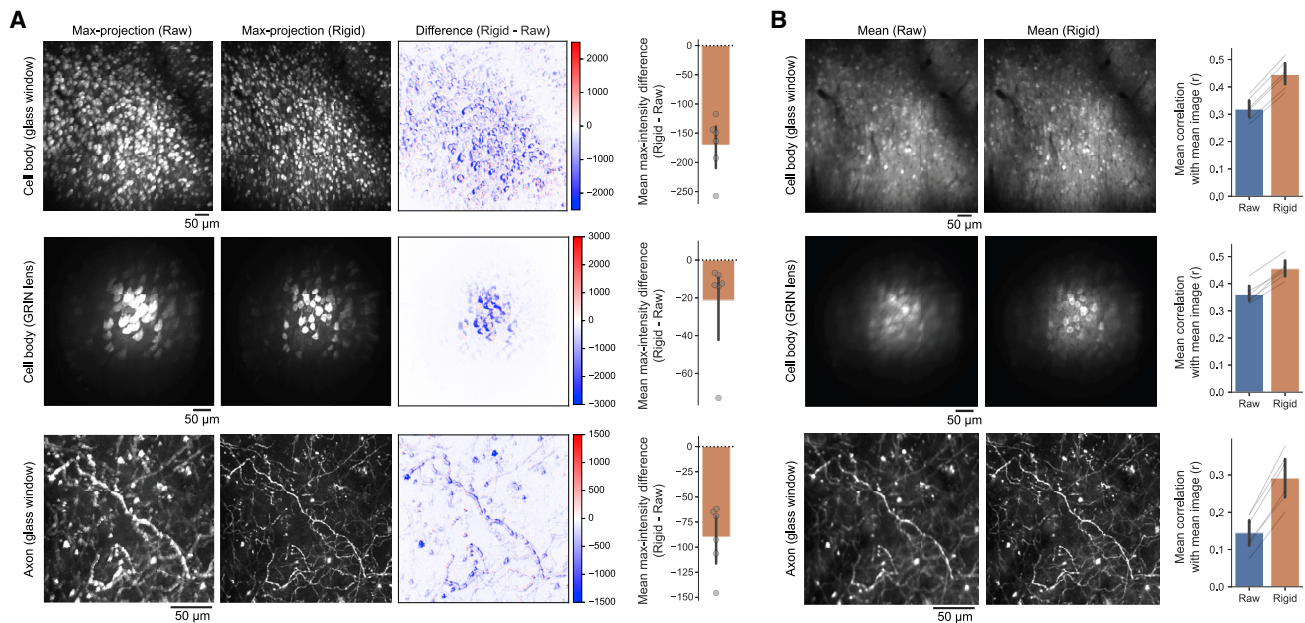
Figure 3. Performance of rigid corrections of motion artifacts

(A) Max-intensity difference between pre- and post-rigid motion corrections. Max-projection images of raw frames, motion-corrected frames, and their intensity difference are shown. The bar graph indicates the mean max-intensity difference (mMD) within each FOV. mMD was negative in all sessions, indicating that motion artifacts consistently decreased after rigid motion corrections. Each circle indicates the result from an imaging session. The error bars are 95% confidence interval (CI).
(B) Mean images of raw frames, motion-corrected frames, and the mean correlations between the mean images and individual frames of each condition (self-mCM). Self-mCM increased in all sessions, indicating that motion artifacts consistently decreased after rigid motion corrections. Each thin gray line indicates the comparison of the same imaging session across conditions. The error bars are 95% CI.
See also Figure S3.

projection image across all frames is calculated for pre- and post-registered images separately. mMD is the difference in the means of all intensity values across pixels between these two images (mMD = post-mean − pre-mean). The mean of max-projection intensities across pixels takes a larger value in the presence of motion artifacts because movement of a neuronal structure across frames increases the total number of pixels to which the neuronal structure with calcium signals contributes to the max-projection image. Therefore, negative mMD indicates reduced motion artifacts after registration. Note that we downsampled frames by 50-frame, non-overlapping moving averaging before creating the max-projection images to suppress the contribution of image noise to the max-projection intensity. The second metric, mCM, is the mean correlation between individual frames and a mean image. mCM increases after successful registration because motion artifacts reduce the similarity between individual frames and the mean image. The mean image with which individual frames are compared for the calculation of correlation coefficient can be the mean of either pre- or post-registered frames. We define mCM as "self-mCM" when image frames are compared with their own mean image and as "cross-mCM" when image frames are compared with the mean image of the other compared condition (e.g., correlation between individual pre-registered frames and the mean image of post-registered frames).

First, we examined the performance of the rigid motion-correction step. In all imaging sessions across the three different

imaging conditions, mMD was consistently negative (Figure 3A). Furthermore, both self-mCM (Figure 3B) and cross-mCM (Figure S3) also consistently improved after registration. Therefore, rigid motion-correction step significantly reduced motion artifacts in all imaging sessions (Video S1).

Next, we examined the performance of the warp-correction step. We compared mMD and mCM metrics between images after the rigid motion-correction step and images after both the rigid motion- and warp-correction steps. We found that mMD ([two-step mean] − [rigid mean]) was consistently negative in all sessions (Figure 4A). Furthermore, both self-mCM (Figure 4B) and cross-mCM (Figure S4) consistently improved in all sessions, and the improvement on the self-mCM was observed across all patches (Figure S5). These results indicate that slow image distortions are ubiquitous in *in vivo* two-photon calcium imaging data from awake-behaving mice, and PatchWarp successfully corrects the distortions (Videos S2 and S3). The consistent improvement across eight different cortical areas (ALM, pM2, RSC, PPC, S1, V1, OFC, and M1) with three different imaging conditions (cell bodies or axons through glass windows and cell bodies through GRIN lenses) indicate the general applicability of PatchWarp algorithm on two-photon calcium data.

We also analyzed the processing time of the PatchWarp pipeline on a standard desktop PC (Intel Core i7-9800X, 32 GB RAM, Windows 10). The processing speed depends on the number of available central processing unit (CPU) cores because the algorithm is parallelized at both rigid- and warp-correction steps. We
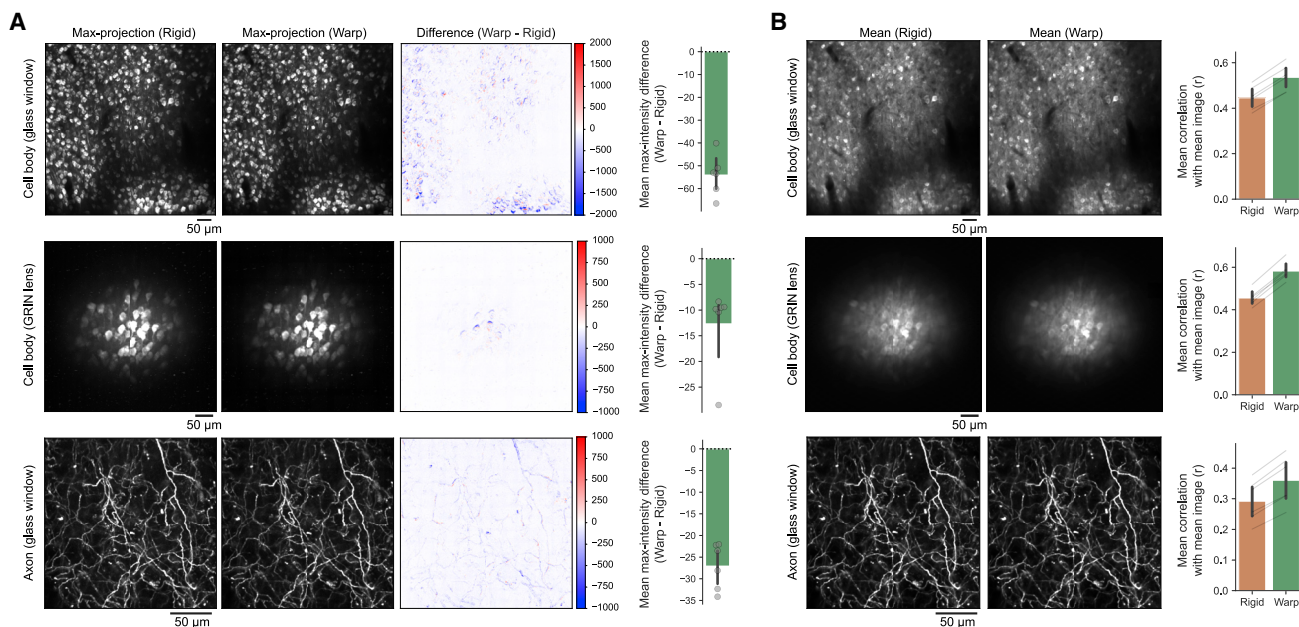
Figure 4. Performance of warp corrections of slow image distortions

(A) Max-intensity difference between images with only rigid motion corrections and images with both rigid and warp corrections. Max-projection images of frames with only rigid motion corrections, frames with both rigid and warp corrections, and their intensity difference are shown. The bar graph indicates the mMD within each FOV. mMD was negative in all sessions, indicating that warp corrections consistently improved the frame-by-frame registrations. Each circle indicates the result from an imaging session. The error bars are 95% CI.

(B) Mean images of frames with only rigid motion corrections, frames with both rigid and warp corrections, and the mean correlations between the mean images and individual frames of each condition (self-mCM). Self-mCM increased in all sessions, indicating that warp corrections consistently improved the frame-by-frame registrations. Each thin gray line indicates the comparison of the same imaging session across conditions. The error bars are 95% CI.

See also Figures S4 and S5.

plot the processing time of six imaging sessions when eight CPU cores are used (Figure 5). The processing time of the rigid correction step was shorter than the imaging time (29-Hz frame rate). Although the processing time was longer than the imaging time when warp corrections were applied, the processing time is still within a reasonable range, even for these exceptionally long imaging sessions. Note that the processing time can be substantially sped up if users have access to computing clusters with many CPU cores.

Performance comparison to existing non-rigid registration algorithms

We compared the performance of PatchWarp with other existing non-rigid image-registration algorithms. These non-rigid registration algorithms for calcium imaging were validated with either two-photon images (Giovannucci et al., 2019; Pachitariu et al., 2017; Pnevmatikakis and Giovannucci, 2017) or one-photon GRIN lens images in the past (Chen et al., 2019; Lu et al., 2018; Table S1). Suite2p (Pachitariu et al., 2017) and CaImAn (Giovannucci et al., 2019) are two of the most popular software for the processing of two-photon calcium imaging data. Therefore, we focus our comparisons on these two algorithms. Both toolboxes offer non-rigid motion-correction functions. The non-rigid motion-correction algorithm in CaImAn is called NoRMCorre (Pnevmatikakis and Giovannucci, 2017), but we call it CaImAn here for simplicity. Non-rigid registration

algorithms in Suite2p and CaImAn split a FOV into patches as in the warp-correction step of PatchWarp. However, both Suite2p and CaImAn estimate only the translational shift of each patch. In contrast, the warp-correction step of PatchWarp estimates affine transformations that can transform each patch by not only translation but also rotation, scaling, and shearing. Therefore, theoretically, PathWarp can correct more complex distortions than Suite2p and CaImAn. Furthermore, Suite2p and CaImAn use cross-correlation or phase correlation to estimate the best translation, while PatchWarp uses ECC. ECC has the property that it is invariant to changes in bias, gain, brightness, and contrast (Evangelidis and Psarakis, 2008). The use of ECC, along with the local intensity normalization (Figure S1), may also improve the registration performance.

We compared the processing time and the registration accuracy of their non-rigid registrations. We used Python versions of Suite2p (v.0.9.0) and CaImAn (v.1.9.3) for this comparison. To make a consistent comparison, we used the same number of patches and the same patch size for each imaging session across the three algorithms, and they were run on the same PC (Intel Core i7-9800X, 32 GB RAM, Windows 10). The processing time of PatchWarp was shortest on 5/6 sessions and was slightly longer than that of Suite2p on one session with an exceptionally large number of frames (379,127 frames; Figure 6A). Furthermore, mMD relative to PatchWarp was positive on most images from Suite2p and CaImAn (Figure 6B). Self-mCM was
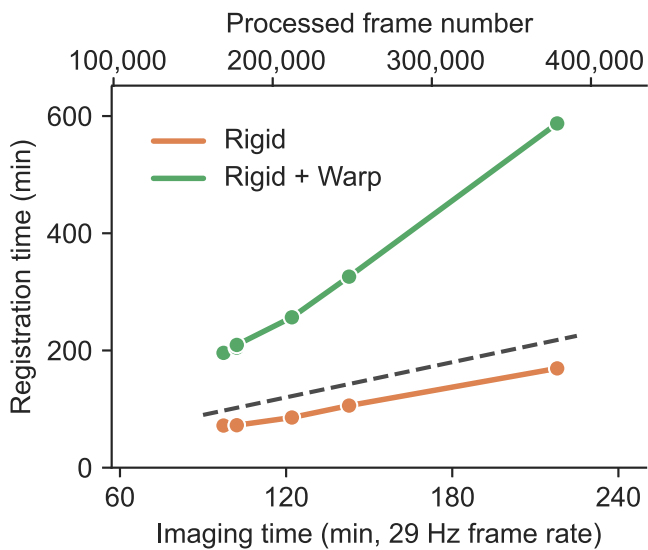
Figure 5. Processing speed of PatchWarp pipeline

Processing time of PatchWarp pipeline on six imaging sessions (cell body imaging, glass window). The pipeline was run on a standard desktop PC. The top and bottom x axes indicate the frame number and the corresponding imaging time from each imaging session. The dashed line represents the unity line for the registration time of the y axis and imaging time of the x axis. The rigid registration takes less time than the imaging time. The registration time exceeds the imaging time when warp corrections are implemented, but the total time is within a reasonable range.

consistently highest with PatchWarp in all imaging sessions (Figure 6C). Thus, at least in the datasets tested here, PatchWarp outperforms the commonly used algorithms for non-rigid motion correction.

PatchWarp algorithm for fully automatic across-session registration

One of the major advantages of two-photon calcium imaging over electrode recordings of neuronal activity is that we can reliably track the identical population of neurons across many days. On each day, experimenters try to find the consistent focal plane by adjusting the position of an objective lens. However, some variability in the position of the focal plane across days is inevitable (e.g., x-y-z coordinate, rotation angle of the FOV, and rotation angle of the objective lens). In particular, the difference in the rotation angle of the objective lens introduces non-uniform image distortion between images from different imaging sessions. Furthermore, the relative positions of neuronal structures in a FOV can also slowly change across days. These non-uniform image distortions necessitate an image-distortion-correction algorithm to optimally register different imaging sessions for longitudinal neural activity analyses. We devised a two-step PatchWarp method to estimate the optimal transformation that can be used to transform images from a session to the coordinate of another session. Unlike the within-session registration that uses individual frames, the inputs to the algorithm for across-session registration are summary images (e.g., mean image and max-projection image) from each imaging session. As the demonstration, we used two RSC imaging sessions that

were acquired on different days and estimated the transformation that optimally transforms the session H image to the session G coordinate (Figures 7A and 7B). We artificially made the registration problem harder than the original condition by adding shifts (translation and rotation) and barrel distortion to demonstrate the robustness of our algorithm, even in the presence of large displacements and complex distortion.

The first step finds the optimal Euclidean transformation that maximizes ECC between the session G summary images and the transformed session H summary images. The whole FOVs are used without splitting them into patches in this first step. The Euclidean transformation is a special case of affine transformation with reduced number of free parameters (six and three parameters in affine and Euclidean transformation matrices, respectively), and the transformation is limited to translation and rotation. Unlike the rigid motion-correction step of the within-session PatchWarp processing, where the corrections were limited to translations, we use Euclidean transformations here because mismatch in the rotation angles of the imaging FOVs is common across different imaging sessions. To ensure the robust estimation, even in the presence of a large displacement between the two imaging sessions, we took the pyramid-based approach by sequentially updating the transformation matrix from the upper-level pyramids at the gradient-based parameter estimation step (see STAR Methods). The transformation that maximizes ECC between the session G images and the transformed session H images was obtained. We locally normalized the intensity of the input summary images before calculating ECC to focus on neuronal structures instead of calcium levels, similarly to the within-session warp-correction method (Figure S1). We use two types of summary images (mean and max-projection images) and obtain two transformation matrices. We select the one of them that results in a higher ECC and apply it to both of the session H summary images.

The second step finds the optimal affine transformations that maximize ECCs between the session G patches and the transformed session H patches. Large displacements are already corrected by the first Euclidean transformation, so the goal of this second set of transformations is to correct the non-uniform image distortions between the two imaging sessions. The affine transformation matrices with six parameters can transform image patches by shearing and scaling in addition to translation and rotation. By selecting the transformation matrix with larger ECC for each patch from the mean and the max-projection image pairs, we can optimally correct non-uniform image distortions, even when a patch lacks sufficient landmarks in one of the summary images.

To quantify the across-session registration performance, we ran the algorithm on six longitudinally imaged neural populations (cell body imaging and glass window). For each population, summary images from a later session were registered to the summary images from the earlier session. The image correlation between early and late sessions consistently improved (Figure 7C). The results also showed that registration performance was better when both rigid and non-rigid corrections were applied.

The two-step PatchWarp algorithm for across-session registration provides us with the transformation matrices that can optimally transform the session H images to the session G
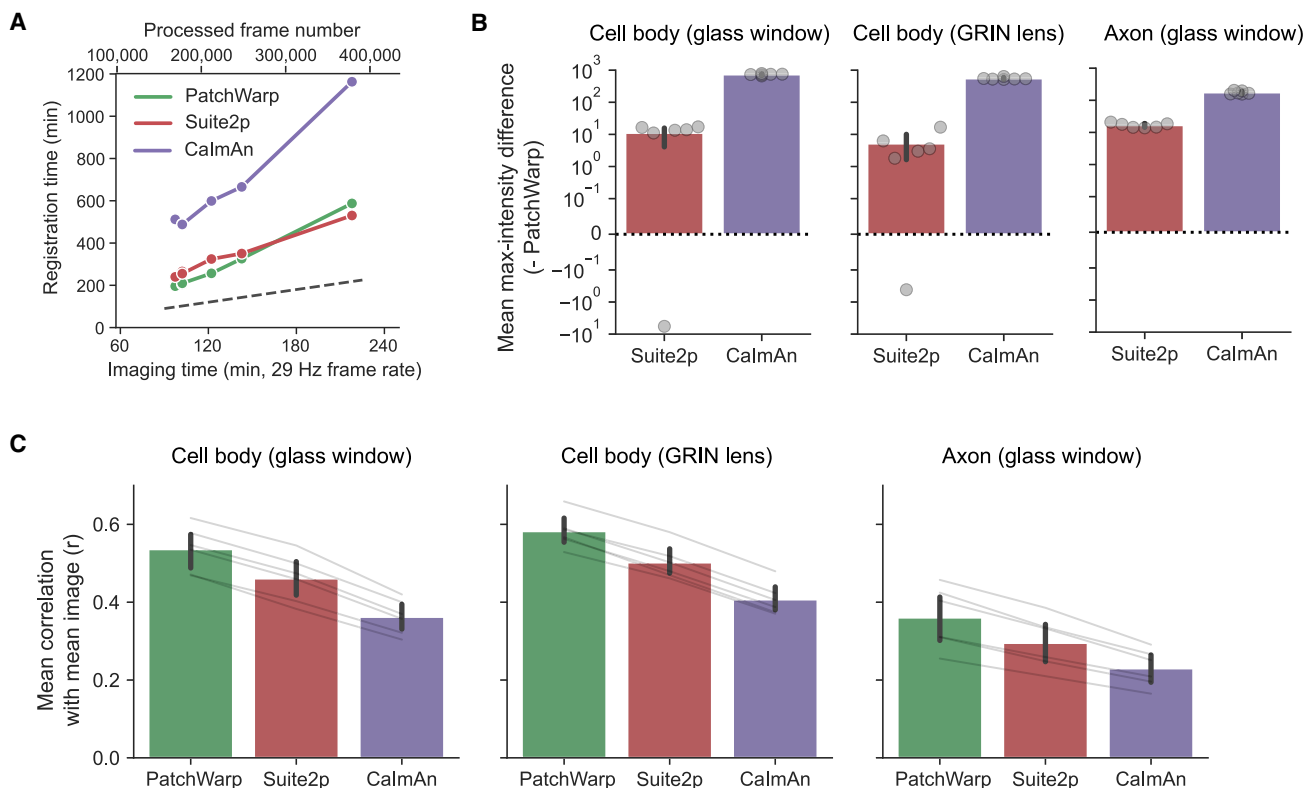
Figure 6. Performance comparison to existing non-rigid registration algorithms

(A) Comparisons of the processing time on the same six imaging sessions (cell body imaging, glass window) run on a standard desktop PC. The top and bottom x axes indicate the frame number and the corresponding imaging time from each imaging session. The dashed line represents the unity line for the registration time of the y axis and imaging time of the x axis.

(B) mMD between images with PatchWarp registrations and images with non-rigid registrations of either Suite2p or CaImAn. mMD was positive in most sessions except for two Suite2p sessions, indicating the generally higher registration accuracy with PatchWarp. Each circle indicates the result from an imaging session. The error bars are 95% CI. The y axis is symmetric log scale where the log scale is applied to both positive and negative values.

(C) Mean correlations between the mean images and individual frames (self-mCM) for images that were processed by PatchWarp or non-rigid registration algorithms of Suite2p and CaImAn. Self-mCM was highest with PatchWarp for all imaging sessions, indicating that registration accuracy was consistently higher with PatchWarp. Each thin gray line indicates the comparison of the same imaging session across conditions. The error bars are 95% CI.

coordinate. The obtained transformation can be applied to any images in the session H coordinate (e.g., individual frames; summary images, such as mean and max projection; correlation map, and ROI mask), and the transformed images can be directly compared with the session G images for longitudinal activity analyses. We provide functions to calculate the transformation and apply the transformation to other images in our software.

DISCUSSION

Removal of activity-independent artifacts from calcium imaging data is critical to accurately interpret data. For example, rigid motion artifacts may be misinterpreted as movement-related neural activity, and slow image distortions may be misinterpreted as slow activity changes over time. These artifacts can have profound impacts on the neural activity analyses and potentially lead to incorrect scientific conclusions. Here, we developed a novel algorithm, PatchWarp, that robustly corrects both the rigid motion artifacts and slow image distortions. We

publicly share the software at https://github.com/ryhattori/PatchWarp for the community.

The warp-correction step of PatchWarp splits a FOV into smaller patches, and optimal affine transformation is estimated for each patch. Two existing and popular motion-correction algorithms for two-photon calcium imaging data, Suite2p and CaImAn (NoRMCorre), offer non-rigid registration algorithms that similarly split a FOV into small patches. However, the transformation of each patch in these published methods is limited to translational shifts. Although translational shifts of many small patches can approximate shearing transformation, the accuracy is suboptimal. The rigidity on each patch limits the capacity of the fine-scale distortion corrections unless each patch is very small, and the patch cannot be too small, given the limited number of spatial landmarks. In contrast, PatchWarp applies an affine transformation to each patch. Affine transformation can transform each patch not only by translation but also rotation, scaling, and shearing. Utilizing the diverse transformational capacity of the affine transformation, we achieved a higher registration accuracy than Suite2p and CaImAn (NoRMCorre) (Figure 6).

Figure 7. PatchWarp algorithm for fully automatic across-session registration

(A) Registration between different imaging sessions for longitudinal analyses (red: session G images; cyan: session H images). Mean- and max-projection images were used as the inputs to the algorithm. The first step finds the optimal Euclidean transformation that corrects large displacements with rigidity (translation and rotation). The second step finds the optimal affine transformations for small patches to correct non-uniform distortions between the two image sessions. The transformations are separately estimated for mean and max-projection images, but only one transformation with larger ECC between G and transformed H, $(\rho_{G,\,E(H)}, \rho_{G_i, A_i\{E(H_i)\}})$, is selected at each step. The obtained transformations can be used to transform any images from session H to the session G coordinate for longitudinal neural activity analyses.

(B) Zoomed images of the subfield that is indicated by yellow dashed lines in (A). This subfield highlights the necessity of the warp-correction step.

(C) Correlations of mean or max-projection images between early and late imaging sessions of the same neural population (six different FOVs, cell body imaging, glass window). The correlation is compared between pre-registration (raw), post-Euclidean transformation (rigid), and post-patchwork affine transformations (warp). Warp correction consistently improved the registration accuracy between different imaging sessions. Each thin gray line indicates the comparison of the same imaging session across conditions. The error bars are 95% CI.

Despite the computational complexity of finding optimal affine transformation, we demonstrated the computational efficiency of our algorithm by parallelizing the registration steps. Although the flexibility of the affine transformation is likely the primary factor that contributed to the high registration accuracy of PatchWarp, other algorithmic differences between different software might have also contributed to the differences in the registration accuracy. For example, the template matching at the non-rigid registration steps use different objective functions between different algorithms (ECC for PatchWarp, phase correlation for Suite2p, and cross-correlation for CaImAn). Further-

more, PatchWarp is unique in applying local intensity normalization that normalizes the temporal signal dynamics (Figure S1).

The warp correction step of PatchWarp relies on a parametric image alignment using ECC (Evangelidis and Psarakis, 2008; Psarakis and Evangelidis, 2005). The ECC algorithm searches for an affine transformation that registers images using a gradient-based method. We adapted the ECC-based algorithm to a two-step piecewise scheme and optimized the image-processing steps such that it stably works on temporally dynamic calcium imaging data. Although a similar piecewise affine transformation method was previously used to register static MRI

images (Batchelor et al., 2005; Pitiot et al., 2003), we showed that PatchWarp reliably works on calcium imaging data, a more challenging dataset with temporally dynamic signals. There are also various diffeomorphic mapping algorithms for non-rigid registrations (Batchelor et al., 2005; Beg and Khan, 2006; Vercauteren et al., 2009), and future studies should test their accuracy and computational efficiency on two-photon calcium imaging data.

We also developed a PatchWarp approach for fully automated across-session image registrations. Registration across images from different imaging sessions is necessary to longitudinally analyze how individual neurons or neuronal structures change activity patterns over days. Since our method uses only the summary images from each session, the estimation of the optimal transformation matrices can be obtained instantly with only 3–6 s on standard desktop computers. By applying the obtained transformations, we can transform images to the coordinate of the other session for longitudinal neural activity analyses.

Limitations of the study

Although PatchWarp is robust in correcting rigid motion artifacts and slow image distortions, we still observed residual motion artifacts in some cases, especially when the neuronal structures are very close to the midline sinus. The residual motion artifacts are due to fast dilation of the midline sinus, which occasionally happens in behaving mice. The dilation of sinus introduces fast image distortions. Since PatchWarp estimates the affine transformations using mean of multiple frames for the robustness and the computational efficiency, it mostly focuses on correcting slow image distortions. Although users can apply PatchWarp to correct fast image distortions by using smaller number of frames to make the mean image to register to the template image, the accuracy of warp correction could be compromised if each mean image does not contain sufficient neuronal structures with visible calcium signals. The dynamic nature of calcium signals introduces uncertainty about the availability of sufficient signals on each mean image. If users need to correct fast image distortions, one potential solution is to co-express stable red fluorescence indicators (e.g., tdTomato; Drobizhev et al., 2011; Shaner et al., 2004) and use the signals for the registration instead of using the dynamic calcium signals. Calcium sensors with high baseline fluorescence, such as jGCaMP7b (Dana et al., 2019) and jGCaMP8 series (Zhang et al., 2021), may also work for robust fast distortion corrections. Alternatively, experimenters should always try to minimize the contributions of the sinus dilations by applying enough pressure when they implant a glass window onto the brain or avoid selecting a FOV near the midline sinus.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animals
- METHOD DETAILS
  - Surgery
  - Two-photon calcium imaging
  - Template re-estimation for rigid motion correction
  - Registration algorithm for rigid motion correction
  - Registration algorithm for slow distortion correction
  - Registration between different imaging sessions
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Data analysis and quantification

AUTHOR CONTRIBUTIONS

R.H. conceived the methods, developed the software, performed analyses, and collected cell body imaging data. R.H. and T.K. wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Adelson, E.H., Anderson, C.H., Bergen, J.R., Burt, P.J., and Ogden, J.M. (1984). Pyramid methods in image processing. RCA Eng. *29*, 33–41.

Batchelor, P.G., Atkinson, D., Irarrazaval, P., Hill, D.L.G., Hajnal, J., and Larkman, D. (2005). Matrix description of general motion correction applied to multishot images. Magn. Reson. Med. *54*, 1273–1280.

Beg, M.F., and Khan, A. (2006). Computing an average anatomical atlas using LDDMM and geodesic shooting. In 2006 3rd IEEE International Symposium on Biomedical Imaging From Nano to Macro - Proceedings (IEEE), pp. 1116–1119.

Broussard, G.J., Liang, Y., Fridman, M., Unger, E.K., Meng, G., Xiao, X., Ji, N., Petreanu, L., and Tian, L. (2018). In vivo measurement of afferent activity with axon-specific calcium imaging. Nat. Neurosci. *21*, 1272–1280.

Chen, T.W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature *499*, 295–300.

Chen, Z., Blair, H.T., and Cong, J. (2019). LANMC: LSTM-assisted non-rigid motion correction on FPGA for calcium image stabilization. In Proceedings

2019 ACM/SIGDA International Symposium Field-Programmable Gate Arrays, pp. 104–109.

Dana, H., Sun, Y., Mohar, B., Hulse, B.K., Kerlin, A.M., Hasseman, J.P., Tsegaye, G., Tsang, A., Wong, A., Patel, R., et al. (2019). High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. Nat. Methods *167*, 649–657.

Drobizhev, M., Makarov, N.S., Tillo, S.E., Hughes, T.E., and Rebane, A. (2011). Two-photon absorption properties of fluorescent proteins. Nat. Methods *8*, 393–399.

Dubbs, A., Guevara, J., and Yuste, R. (2016). Moco: fast motion correction for calcium imaging. Front. Neuroinform. *0*, 6.

Evangelidis, G.D., and Psarakis, E.Z. (2008). Parametric image alignment using enhanced correlation coefficient maximization. IEEE Trans. Pattern Anal. Mach. Intell. *30*, 1858–1865.

Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B.L., Koay, S.A., Taxidis, J., Najafi, F., Gauthier, J.L., Zhou, P., et al. (2019). CalmAn an open source tool for scalable calcium imaging data analysis. Elife *8*, e38173.

Grienberger, C., and Konnerth, A. (2012). Imaging calcium in neurons. Neuron *73*, 862–885.

Guizar-Sicairos, M., Thurman, S.T., and Fienup, J.R. (2008). Efficient subpixel image registration algorithms. Opt. Lett. *33*, 156–158.

Hattori, R., and Komiyama, T. (2022). Context-dependent persistency as a coding mechanism for robust and widely distributed value coding. Neuron *110*, 502–515.e11.

Hattori, R., Kuchibhotla, K.V., Froemke, R.C., and Komiyama, T. (2017). Functions and dysfunctions of neocortical inhibitory neuron subtypes. Nat. Neurosci. *20*, 1199–1208.

Hattori, R., Danskin, B., Babic, Z., Mlynaryk, N., and Komiyama, T. (2019). Area-specificity and plasticity of history-dependent value coding during learning. Cell *177*, 1858–1872.e15.

Hunter, J.D. (2007). Matplotlib: a 2D graphics environment. Comput. Sci. Eng. *9*, 90–95.

Inoue, M., Takeuchi, A., Manita, S., Horigane, S., Sakamoto, M., Kawakami, R., Yamaguchi, K., Otomo, K., Yokoyama, H., Kim, R., et al. (2019). Rational engineering of XCaMPs, a multicolor GECI suite for in vivo imaging of complex brain circuit dynamics. Cell *177*, 1346–1360.e24.

Lu, J., Li, C., Singh-Alvarado, J., Zhou, Z.C., Fröhlich, F., Mooney, R., and Wang, F. (2018). MIN1PIPE: a miniscope 1-photon-based calcium imaging signal extraction pipeline. Cell Rep. *23*, 3673–3684.

Mayford, M., Bach, M.E., Huang, Y.Y., Wang, L., Hawkins, R.D., and Kandel, E.R. (1996). Control of memory formation through regulated expression of a CaMKII transgene. Science *274*, 1678–1683.

Mitani, A., and Komiyama, T. (2018). Real-time processing of two-photon calcium imaging data including lateral motion artifact correction. Front. Neuroinform. *12*, 98.

Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L.F., Dalgleish, H., Carandini, M., and Harris, K.D. (2017). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. Preprint at bioRxiv. https://doi.org/10.1101/061507.

Pitiot, A., Malandain, G., Bardinet, E., and Thompson, P.M. (2003). Piecewise affine registration of biological images. Int. Work. Biomed. Image Regist., 91–101.

Pnevmatikakis, E.A., and Giovannucci, A. (2017). NoRMCorre: an online algorithm for piecewise rigid motion correction of calcium imaging data. J. Neurosci. Methods *291*, 83–94.

Psarakis, E.Z., and Evangelidis, G.D. (2005). An enhanced correlation-based method for stereo correspondence with sub-pixel accuracy. Proc. IEEE Int. Conf. Comput. Vis., 907–912.

Rossi, J., Balthasar, N., Olson, D., Scott, M., Berglund, E., Lee, C.E., Choi, M.J., Lauzon, D., Lowell, B.B., and Elmquist, J.K. (2011). Melanocortin-4 receptors expressed by cholinergic neurons regulate energy balance and glucose homeostasis. Cell Metab. *13*, 195–204.

Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N.G., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat. Biotechnol. *22*, 1567–1572.

Svoboda, K., and Yasuda, R. (2006). Principles of two-photon excitation microscopy and its applications to neuroscience. Neuron *50*, 823–839.

Thévenaz, P., Ruttimann, U.E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. IEEE Trans. Image Process. *7*, 27–41.

Vercauteren, T., Pennec, X., Perchant, A., and Ayache, N. (2009). Diffeomorphic demons: efficient non-parametric image registration. Neuroimage *45*, S61–S72.

Waskom, M.L. (2021). seaborn: statistical data visualization. J. Open Source Softw. *6*, 3021.

Wekselblatt, J.B., Flister, E.D., Piscopo, D.M., and Niell, C.M. (2016). Large-scale imaging of cortical dynamics during sensory perception and behavior. J. Neurophysiol. *115*, 2852–2866.

Zhang, Y., Rózsa, M., Liang, Y., Bushey, D., Wei, Z., Zheng, J., Reep, D., Broussard, G.J., Tsang, A., Tsegaye, G., et al. (2021). Fast and sensitive GCaMP calcium indicators for imaging neural populations. Preprint at bioRxiv. https://doi.org/10.1101/2021.11.08.467793.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
AAV-hSyn1-FLEx-axon-GCaMP6s	Broussard et al., 2018	Addgene: 112010-AAV5
Deposited data		
Example imaging dataset for motion and distortion correction	This paper	https://doi.org/10.6084/m9.figshare.19212501.v1
Example imaging dataset for across-session registration	This paper	https://doi.org/10.6084/m9.figshare.19217421.v1
Experimental models: Organisms/strains		
Mouse: CaMKIIa-tTA: B6;CBA-Tg(Camk2a-tTA)1Mmay/J	The Jackson Laboratory	RRID:IMSR_JAX:003010
Mouse: tetO-GCaMP6s: B6;DBA-Tg(tetO-GCaMP6s)2Niell/J	The Jackson Laboratory	RRID:IMSR_JAX:024742
Mouse: B6;129S6-Chattm2(cre)Lowl/J	The Jackson Laboratory	RRID:IMSR_JAX:006410
Software and algorithms		
PatchWarp	This paper	https://doi.org/10.5281/zenodo.5232757, https://github.com/ryhattori/PatchWarp
MATLAB	MathWorks	RRID:SCR_001622
Python3	Python Software Foundation	RRID:SCR_008394
ScanImage	Vidrio Technologies, LLC	RRID:SCR_014307
suite2p	Pachitariu et al., 2017, bioRxiv	https://github.com/MouseLand/suite2p
CaImAn	Giovannucci et al., 2019, eLife	RRID:SCR_021533 https://github.com/flatironinstitute/CaImAn
Matplotlib	Hunter, Computing in Science and Engineering, 2010, pp. 90-95, vol. 9	https://github.com/matplotlib/matplotlib
Seaborn	Waskom, 2021, Open Source Software, 6, 3021	https://github.com/mwaskom/seaborn

RESOURCE AVAILABILITY

Lead contact
Further information and requests for code should be directed to and will be fulfilled by the lead contact, Ryoma Hattori (rhattori0204@gmail.com).

Materials availability
This study did not generate new unique reagents.

Data and code availability
- Example image datasets to run the PatchWarp software are available for within-session registration (https://doi.org/10.6084/m9.figshare.19212501.v1) and across-session registration (https://doi.org/10.6084/m9.figshare.19217421.v1).
- PatchWarp software (https://doi.org/10.5281/zenodo.5232757) is available at https://github.com/ryhattori/PatchWarp.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals
Mice were originally obtained from the Jackson Laboratory (CaMKIIa-tTA: B6;CBA-Tg(Camk2a-tTA)1Mmay/J [JAX 003010]; tetO-GCaMP6s: B6;DBA-Tg(tetO-GCaMP6s)2Niell/J [JAX 024742]; B6;129S6-Chattm2(cre)Lowl/J [JAX:006410]). All animals were

housed in disposable plastic cages with standard bedding in a room on a reversed light cycle (12 h/12 h). All procedures were performed following protocols approved by the UCSD Institutional Animal Care and Use Committee (IACUC) and guidelines of the National Institute of Health.

METHOD DETAILS

Surgery

Surgical procedures were performed as previously described (Hattori and Komiyama, 2022; Hattori et al., 2019). Mice were continuously anesthetized with 1%–2% isoflurane during surgery. We subcutaneously injected dexamethasone (2mg/kg). We exposed the dorsal skull, removed the connective tissue on the skull surface using a razor blade, and performed craniotomy. For axon imaging experiments, ~1 μL of AAV-hSyn-FLEX-axon-GCaMP6s virus (Broussard et al., 2018) was unilaterally injected in the basal forebrain of ChAT-Cre mice (~0.3 mm posterior and ~1.7 mm lateral from bregma, 5.0 mm deep from the brain surface). After the injection, a glass window was secured on the edges of the remaining dorsal skull using 3M Vetbond (WPI), followed by cyanoacrylate glue and dental acrylic cement (Lang Dental). For GRIN lens imaging experiments, we first aspirated the cortex above the target coordinate up to 1.0 mm depth using a blunt end 30G needle (0.312 mm O.D., SAI Infusion Technologies). Then, a GRIN lens (500 μm diameter; Inscopix, GLP-0561) was unilaterally implanted above the deep layer of lateral OFC ((~2.45 mm lateral and ~2.6 mm anterior from bregma, 1.5 mm deep from the brain surface). The implanted GRIN lens was secured using 3 M Vetbond (WPI) on the skull, followed by cyanoacrylate glue and dental acrylic cement (Lang Dental). After the implantation of either a glass window or a GRIN lens, a custom-built metal head-bar was secured on the skull above the cerebellum using cyanoacrylate glue and dental cement. Mice were subcutaneously injected with Buprenorphine (0.1 mg/kg) and Baytril (10 mg/kg) after surgery.

Two-photon calcium imaging

Neural calcium signals were recorded using two-photon microscopes (B-SCOPE, Thorlabs) with a 16×, 0.8 NA water immersion objective lens (Nikon) and 925 nm lasers (Ti-Sapphire laser, Newport) from head-fixed behaving mice. ScanImage (Vidrio Technologies) running on MATLAB (MathWorks) was used for image acquisitions. Images (512 × 512 pixels) were continuously recorded at ~29 Hz during each imaging session. Cell body imaging sessions were recorded for 1.5–3.5 h while axon imaging sessions were recorded for ~20 min. Imaging was performed at superficial layers of cortex for glass window imaging, and GRIN lens imaging was performed at deep layers of OFC. The central coordinates of FOVs from bregma were 1.7 mm lateral and 2.25 mm anterior for ALM, 0.4 mm lateral and 0.5 mm anterior for pM2, 0.4 mm lateral and 2 mm posterior for RSC, 1.7 mm lateral and 2 mm posterior for PPC, 1.8 mm lateral and 0.75 mm posterior for S1, 2.5 mm lateral and 3.25 mm posterior for V1.

Template re-estimation for rigid motion correction

Rigid motion artifacts during calcium imaging were corrected by registering individual frames to a template image. The template image was iteratively updated during the motion correction process. These iterative re-estimations of the template image improve the robustness of motion correction performance, especially in the presence of slow image distortion over time. First, we created the 1st template image (T_1) by averaging the middle 2,500 frames of an imaging session. Then, we registered the 2,500 frames to the template T_1 and updated the T_1 by averaging the motion-corrected 2,500 frames (T_1'). Each imaging session was temporally split into 5 blocks, and T_1' was used as the template image to register frames in the 3rd block. After registering all frames in the 3rd block, we created 2 new template images by averaging the first and the last 2,500 motion-corrected frames in the 3rd block (T_2' and T_3', respectively). The template T_2' was used to register all frames in the 2nd block, and the template T_3' was used to register all frames in the 4th block. Again, we created new template images by averaging the first 2,500 frames of the 2nd block (T_4') and averaging the last 2,500 frames of the 4th block (T_5'). The template image T_4' was used to register frames in the 1st block, and the template image T_5' was used to register frames in the 5th block.

Registration algorithm for rigid motion correction

Individual frames were shifted along x and y directions such that the correlation between each frame and a template image is maximized. To estimate the best translational shift of each frame, pyramid method (Adelson et al., 1984) and hill climbing algorithm were combined as reported previously (Mitani and Komiyama, 2018). First, both the template image and individual frames were downscaled to the level 3 pyramid representation (1/8 downscaling factor). Hill climbing algorithm finds the best translation that maximizes the correlation between each frame and the template. Then, the algorithm moves to the level 2 pyramid representation (1/4 downscaling factor) and uses the translation from the level 3 as the initial estimate of the translational shift to obtain the best shift at the level 2. It repeats the same procedure at the level 1 pyramid representation (1/2 downscaling factor) and lastly at the level 0 (original image resolution). The registration was performed in subpixel accuracy by fitting a quadratic function and interpolation. The pixel intensities of both the template image and individual frames were rank transformed before running the pyramid-based hill climbing algorithm.

Registration algorithm for slow distortion correction

We performed affine transformations of images to correct slow image distortion. However, image distortions in calcium imaging data are usually not uniform across the FOV, making it impossible to correct the distortion by a single affine transformation matrix.

To correct non-uniform distortions, we developed a patchwork approach of affine transformation. Our approach assigns different affine transformation matrices to different subfields of an image, and the transformed subfields are stitched together. This patchwork approach efficiently and robustly corrects non-uniform image distortions in calcium imaging data.

We split each FOV of the size H × W (pixels) into M × M square subfields (patches). We used M = 8 for cellular resolution imaging and used M = 12 or 15 for axon imaging. We used larger M for axon imaging because axons were more dynamic and the movements were independent from each other. Each patch had $0.3 \times \frac{H}{M}$ overlapping pixels with its vertically adjacent patches and $0.3 \times \frac{W}{M}$ overlapping pixels with its horizontally adjacent patches. After affine transformation of each patch, all patches were stitched together by averaging the intensity of the overlapping pixels.

The template image for each patch was created by averaging the middle 5,500 frames of an imaging session. We then temporally downsampled all frames by non-overlapping moving averaging of 500 frames. Optimal affine transformation for each patch was obtained for each of the downsampled frames. This downsampling significantly decreases the computation time and did not affect our results because the slow image distortions were negligible within 500 frames (~17 sec, ~29 Hz). Modern 2-photon microscopy scans a FOV using a resonant scanner with high-frequency scanning rate (~10k Hz), leading to a high frame rate. At high frame rates (>10 Hz), motion artifacts induced by body movements of behaving animals are mostly image translations with minimal distortions. One exceptional case where fast distortion occurs is when the shape of the imaged brain tissue changes during the motion (e.g. by dilation of midline sinus) as we discussed in the discussion section. Such potential fast distortions can typically be prevented at the surgery step for 2-photon microscopy by applying a pressure on the imaging area when the glass window is implanted.

An affine transformation matrix for a 2D image is given by the following 3 × 3 matrix with 6 parameters;

$$\boldsymbol{A} = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{pmatrix}$$ (Equation 1)

The relationships between the original coordinate (x, y) and the transformed coordinate (x', y') is given by

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$ (Equation 2)

To find the optimal affine transformation for correcting image distortion of each patch, we used a gradient-based algorithm as described previously (Evangelidis and Psarakis, 2008). Our objective is to minimize the difference between a template image patch and a warped image patch, but we want to focus on the geometric difference while ignoring brightness difference because brightness changes across frames on calcium imaging data. We can define the loss function of the gradient-based algorithm as follows;

$$L(\boldsymbol{a}) = \left\| \frac{\overline{\boldsymbol{p_t}}}{\|\boldsymbol{p_t}\|} - \frac{\overline{\boldsymbol{p_w}(\boldsymbol{a})}}{\|\boldsymbol{p_w}(\boldsymbol{a})\|} \right\|^2$$ (Equation 3)

where $\| \ \|$ denotes L2 norm, $\boldsymbol{a} = [a_1, a_2, a_3, a_4, a_5, a_6]$ is a vector of parameters for the affine transformation, $\boldsymbol{p_t}$ is a vector with intensity of all pixels in the template patch, $\boldsymbol{p_w}(\boldsymbol{a})$ is a vector with intensity of all pixels in the patch that was warped by an affine transformation matrix with the parameters $\boldsymbol{a}$. $\overline{\boldsymbol{p_t}}$ and $\overline{\boldsymbol{p_w}(\boldsymbol{a})}$ denote the zero-centered vectors of $\boldsymbol{p_t}$ and $\boldsymbol{p_w}(\boldsymbol{a})$, respectively. The pixel intensity normalizations of both the template and warped images in [Equation 3] ensure that the loss function is invariant to changes in bias, gain, brightness, and contrast. Thus, we can selectively minimize geometric difference using the loss function. We can equivalently express the [Equation 3] as follows;

$$L(\boldsymbol{a}) = 2 - 2 * \frac{\overline{\boldsymbol{p_t}} \cdot \overline{\boldsymbol{p_w}(\boldsymbol{a})}}{\|\boldsymbol{p_t}\| \|\boldsymbol{p_w}(\boldsymbol{a})\|}$$ (Equation 4)

Therefore, minimizing $L(\boldsymbol{a})$ is equivalent to maximizing the following index;

$$\rho(\boldsymbol{a}) = \frac{\overline{\boldsymbol{p_t}} \cdot \overline{\boldsymbol{p_w}(\boldsymbol{a})}}{\|\boldsymbol{p_t}\| \|\boldsymbol{p_w}(\boldsymbol{a})\|}$$ (Equation 5)

where $\overline{\boldsymbol{p_t}}$ and $\overline{\boldsymbol{p_w}(\boldsymbol{a})}$ are zero-centered vectors of $\boldsymbol{p_t}$ and $\boldsymbol{p_w}(\boldsymbol{a})$. This index is called as enhanced correlation coefficient (ECC). We estimated the optimal affine transformation matrix that maximizes ECC using a gradient-based approach (Evangelidis and Psarakis, 2008; Psarakis and Evangelidis, 2005).

To further improve the robustness and accuracy of the gradient-based algorithm, we processed both input images to normalize the temporal dynamics of calcium signals. First, we obtained blurred images by performing 2D convolution using a circular kernel (32 pixel radius) as follows;

$$g(x, y)_{r32} = \omega_{r32} * f(x, y)$$ (Equation 6)

where $*$ denotes convolution operation, $g(x, y)$ is the original image, ω_{r32} is a circular kernel with 32 pixel radius, and $g(x, y)_{r32}$ is the blurred image. The pixel intensity near the edges of the blurred images are low due to zero-padding. To compensate for the zero-padding effect, we normalized the blurred images as follows;

$$g(x, y)'_{r32} = \frac{g(x, y)_{r32}}{\omega_{r32} * J_{H,W}}$$

(Equation 7)

where $J_{H, W}$ is an all-ones matrix where every element is equal to 1. Using the blurred images, we normalized the intensity of the original image as follows;

$$f(x, y)' = \frac{\sum_{x=1}^{W} \sum_{y=1}^{H} f(x, y)^{H}}{H \times W} \times \frac{f(x, y)}{g(x, y)'_{r32}}$$

(Equation 8)

This intensity normalization enhances the contrast between dim neuronal structures and their local surrounding areas (radius of 32 pixels) on each frame. This normalization suppresses the changes in neuronal calcium-dependent fluorescence intensity across frames. The time-invariance of the normalized images helps the gradient-based algorithm to match the target patch to the template patch.

After deriving the optimal affine transformation matrix for each image patch, we transformed each patch by subpixel bilinear interpolation such that the x-y coordinate is transformed following the optimal affine transformation. The transformed patches were stitched together by averaging the intensity of overlapping pixels from adjacent patches to obtain the transformed full-FOV image.

Registration between different imaging sessions

We developed a robust automated approach to register images between different imaging sessions for longitudinal neural activity analyses. The inputs to the algorithm are summary images from each imaging session. It accepts any summary images such as the mean image (intensity is averaged across frames for each pixel), the max-projection image (max-projection of intensity from all frames for each pixel), the standard deviation image (standard deviation of intensity across frames for each pixel), and the correlation image (average correlation across frames between each pixel and its neighbors). Here, we use mean and max-projection images as the example inputs. We denote the mean and max-projection images of session G and session H as $f(x, y)_{mean, G}$, $f(x, y)_{mean, H}$, $f(x, y)_{max, G}$, and $f(x, y)_{max, H}$, respectively. First, we locally normalize the intensity of each summary image using [Equation 8] and obtain $f(x, y)'_{mean, G}$, $f(x, y)'_{mean, H}$, $f(x, y)'_{max, G}$, and $f(x, y)'_{max, H}$. Our goal is to obtain the optimal transformation matrices for the session H summary images such that transformed summary images of session H match the summary images of session G. We obtain the optimal transformation for the session H with a 2-step process.

The first step finds the optimal Euclidean transformation that maximizes ECC between the session G summary images and the transformed session H summary images. The whole FOVs are used without splitting them into patches in this first step. The Euclidean transformation matrix for a 2D image is given by the following 3 × 3 matrix with 3 parameters;

$$\mathbf{E} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & e_1 \\ \sin(\theta) & \cos(\theta) & e_2 \\ 0 & 0 & 1 \end{pmatrix}$$

(Equation 9)

The Euclidean transformation matrix is a special case of an affine transformation matrix, and the transformation is limited to translation (e_1 for x-shift, e_2 for y-shift) and rotation (θ). The reduced number of parameters (6 parameters in $\mathbf{A}$, 3 parameters in $\mathbf{E}$) makes the estimation of the best translational and rotational shifts more robust than when the affine transformation matrix with full 6 parameters are used.

To ensure the robust estimation of $\mathbf{E}$ even in the presence of a large displacement between the 2 imaging sessions, we took the pyramid-based approach. First, intensity-normalized images $f(x, y)'_{mean, G}$ and $f(x, y)'_{mean, H}$ were downscaled to the level 3 pyramid representation (1/8 downscaling factor). A gradient-based algorithm finds the best $\mathbf{E}$ that maximizes ECC [Equation 5] between the session G image $f(x, y)'_{mean, G}$ and the transformed session H image $\mathbf{E}(f(x, y)'_{mean, G})$ using the level 3 pyramid representations. Then, the algorithm moves to the level 2 pyramid representation (1/4 downscaling factor) and uses $\mathbf{E}$ from the level 3 as the initial estimate to obtain the best $\mathbf{E}$ at the level 2. It repeats the same procedure at the level 1 pyramid representation (1/2 downscaling factor) and lastly at the level 0 (original image resolution). We performed the pyramid-based approach for both the mean and max-projection image pairs to obtain the optimal Euclidean transformation matrices $\mathbf{E}_{mean}$ and $\mathbf{E}_{max}$. We then select the Euclidean transformation matrix that gives the largest ECC as follows;

$$\mathbf{E}_{best} = \underset{\mathbf{E}}{\mathrm{argmax}} \left(\rho_{G, \mathbf{E}_k(H)} \right)$$

(Equation 10)

where $\mathbf{E}_k$ is either $\mathbf{E}_{mean}$ or $\mathbf{E}_{max}$, and $\rho_{G, \mathbf{E}_k(H)}$ denotes ECC between the summary image from the session G and the transformed summary image from the session H. We used $\mathbf{E}_{best}$ to transform $f(x, y)'_{mean, H}$ and $f(x, y)'_{max, H}$ to $\mathbf{E}_{best}f((x, y)'_{mean, H})$ and $\mathbf{E}_{best}f((x, y)'_{max, H})$, respectively, by subpixel bilinear interpolation.

The above Euclidean transformation registers the 2 imaging sessions using translational and rotational shifts. However, these rigid shifts are not sufficient for optimal registration due to slight differences in focal planes (e.g. rotation angle of the objective lens,

z-depth of the focal plane) and deformation of the imaged brain tissue between the 2 imaging sessions. Therefore, the next second step aims at solving this non-rigid registration problem using the method we described in the previous section ("*Registration algorithm for slow distortion correction*"). We split $f(x, y)'_{mean, G}$ and $\boldsymbol{E}_{best}f((x, y)'_{mean, H})$ into M × M square patches and find the optimal affine transformation for each patch. We use affine transformation matrices with full 6 parameters [Equation 1] to allow not only translation and rotation but also scaling and shearing for the registration. We do not need to implement pyramid method in this second step because the first Euclidean transformation already corrected large displacements.

We estimated the optimal affine transformations for all patches of both the mean and max-projection image pairs. We denote the optimal affine transformation matrices for i^{th} patch by $\boldsymbol{A}_{i, mean}$ and $\boldsymbol{A}_{i, max}$. We then select the affine transformation matrix that gives the largest ECC for each patch as follows;

$$\boldsymbol{A}_{i,best} = \underset{\boldsymbol{A}_i}{\mathrm{argmax}}\left(\rho_{G_i \boldsymbol{A}_{i,k}\{\boldsymbol{E}_{best}(H_i)\}}\right) \qquad \text{(Equation 11)}$$

where G_i is the ith patch from session G, H_i is the ith patch from session H, $\boldsymbol{A}_{i, k}$ is either $\boldsymbol{A}_{i, mean}$ or $\boldsymbol{A}_{i, max}$, and $\rho_{G_i \boldsymbol{A}_{i,k}\{\boldsymbol{E}_{best}(H_i)\}}$ denotes ECC between the ith patch from the session G and the transformed i^{th} patch from the session H. We can use $\boldsymbol{E}_{best}$ and $\boldsymbol{A}_{i, best}$ to transform i^{th} patches $f(x, y)'_{mean, H_i}$ and $f(x, y)'_{max, H_i}$ to $\boldsymbol{A}_{i,best}\{\boldsymbol{E}_{best}(f(x, y)'_{mean, H_i})\}$ and $\boldsymbol{A}_{i,best}\{\boldsymbol{E}_{best}(f(x, y)'_{max, H_i})\}$, respectively, by subpixel bilinear interpolation. We repeat these steps for all patches, so the number of affine transformation matrices that we obtain is M × M. In total, we obtain (M × M + 1) transformation matrices (including both $\boldsymbol{E}_{best}$ and $\boldsymbol{A}_{i, best}$) for a pair of imaging sessions which we want to register. Using these transformation matrices, we can transform individual imaging frames or other summary images (e.g. mean image, max-projection image, correlation image, standard deviation image, ROI mask image) from session H to session G coordinate for longitudinal neural activity analyses.

Although changes of neural activity across sessions are largely removed by the local intensity normalization [Equation 8], some neurons can be completely invisible on some imaging sessions for unknown reasons that may include inactivity, FOV mismatch, and cell death. PatchWarp was robust as long as each patch contains some shared landmarks between the template and the target images, but the algorithm can fail if there are few shared landmarks between images. In such rare cases, PatchWarp simply keeps the patch as is after the rigid registration step.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data analysis and quantification
PatchWarp software was written in MATLAB. We used MATLAB and Python3.7 for data analyses and visualizations. Matplotlib (Hunter, 2007) and seaborn (Waskom, 2021) were used to create figure plots. All bar plots represent the means of the distributions, and all error bars are 95% confidence intervals (CI). The original codes for pyramid-based hill climbing algorithm (Mitani and Komiyama, 2018) and gradient-based ECC maximization algorithm (Evangelidis and Psarakis, 2008) are available at https://github.com/amitani/matlab_motion_correct and https://www.mathworks.com/matlabcentral/fileexchange/27253, respectively.