Software for Non-Parametric Image Registration of 2-Photon Imaging Data

Philipp Flotho*[1,2], Shinobu Nomura[3], Bernd Kuhn[3], Daniel J. Strauss[1]

[1]**Systems Neuroscience and Neurotechnology Unit**, Neurocenter, Faculty of Medicine, Saarland University & School of Engineering, htw saar, Germany

[2] **Summer Program**, Japan Society for the Promotion of Science (JSPS), Tokyo, Japan

[3]**Optical Neuroimaging Unit**, Okinawa Institute of Science and Technology Graduate University, Tancha, Onna-son, Kunigami, Okinawa, Japan

Abstract

Functional 2-photon microscopy is a key technology for imaging neuronal activity which can, however, contain non-rigid movement artifacts. Despite the established performance of variational *optical flow* (OF) estimation in different computer vision areas and the importance of movement correction for 2-photon applications, no OF-based method for 2-photon imaging is available. We developed the easy-to-use toolbox *Flow-Registration* that outperforms previous alignment tools and allows to align and reconstruct even low *signal-to-noise* 2-photon imaging data.

2-photon microscopy in combination with synthetic or genetically encoded indicators allows to image a wide range of different aspects of neuronal activity with cellular or even sub-cellular resolution in anesthetized as well as behaving animals [1], [2]. Importantly, small signal changes might carry crucial information. However, the imaging data can be afflicted with different types of noise and artifacts. On one hand, due to the low number of generated photons with 2-photon excitation, the shot noise is typically significant. On the other hand, movement noise can be introduced during acquisition. Motion artifacts can be caused by heart beat, breathing, as well as motor behavior in awake animals. Also, some experimental paradigms inherently result in large, non-rigid and non-elastic deformations, for example, local drug injections. While many established tools exist for the correction of small vibrations and rigid drifts, the compensation of large and/or non-uniform motion is still a challenge.

Furthermore, for high accuracy alignment, subpixel registration is necessary. Small displacements can be approximated by a linearized motion assumption [3]. This means even small residual motion can induce large artifacts around image edges due to a proportional relation of spatial image gradient and motion magnitude with respect to induced brightness changes.

While *optical flow* (OF) based image registration methods were used for 2-photon imaging data before [4], they do not incorporate the advances in OF estimation developed in recent years in the context of many computer vision areas.

As a consequence, they perform poorly, when compared with state-of-the-art image registration methods for 2-photon imaging and are generally regarded as too prone to noise for this application [5].

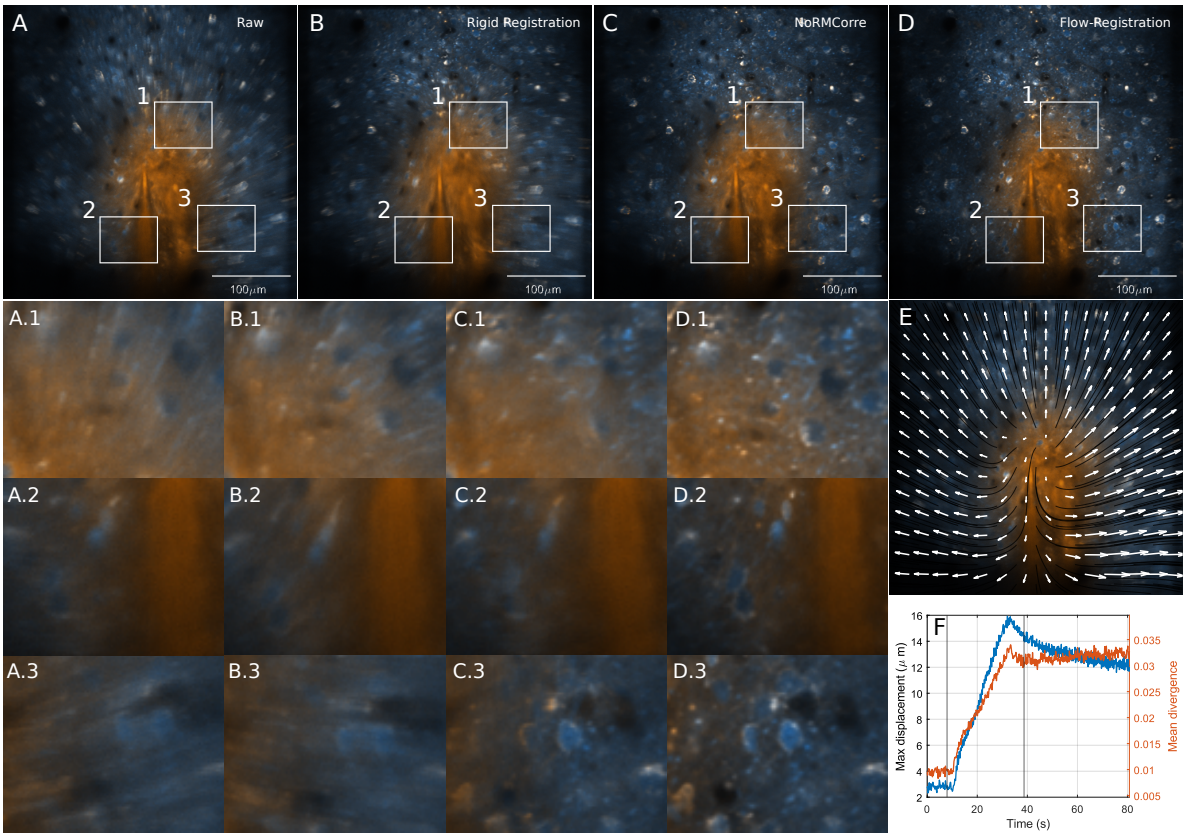

Figure 1: Comparison of registration performance on a very challenging two-channel 2-photon recording during drug injection in vivo. The challenges of the sequence are a very low signal to noise ratio together with brightness changes in the functional imaging channel (orange) and non-elastic deformations due to the injected indicator. Average of (A) raw images, (B) after rigid registration, (C) after registration with NoRMCorre, and (D) after Flow-Registration. The tissue expands from the injection point (E) resulting in large displacements as well as in a high divergence (F) in the displacement field. Flow-Registration can recover fine structures with much more detail, allowing region-of-interest selection along fine structures, while the blur and double images in (A-C) indicate residual motion. While NoRMCorre manages to register the top and middle left image area well, the high divergence bottom half shows residual movement artifacts. The images are a high contrast, false color representations. The channels have all been normalized with respect to the min and max intensity values of the average raw recording.

There has been a great deal of work on OF techniques in the past decades with the goal of improving accuracy, model invariants as well as robustness and computing speed. The recent advancements in OF estimation tackle the problem of large displacements with discontinuous motion fields and motion layers. Chen et al. [6] lead (as of July 2021) the Middlebury optical flow benchmark [7] with respect to *average endpoint error* (AEE) and *average angular error* (AAE). They use similarity transformations for a segmented flow field as initialization of large motion. In a second step, the variational method of Sun et al. [8] is used for subpixel refinement.

In this work, we implement an OF based image registration approach. We build on the well studied framework for variational OF estimation [3], [9], [8]. We adapt this framework to 2-photon imaging data by techniques which have recently been developed in visual computing. We demonstrate the performance on challenging 2-photon imaging data and can report state-of-the-art results in terms of registration quality, competing computation speed and easy accessibility. Our method is available as an easy-to-use MATLAB toolbox as well as an ImageJ plugin.

The motion statistics for 2-photon imaging differ from the motion we encounter in natural images in several ways: Due to the illumination strategy, there is only a single imaging plane and thus no different motion layers. Also, the imaged object is usually soft, biological tissue and as a consequence, we expect smooth motion fields without discontinuities. The image usually contains a fixed region of interest with small displacements between frames and, potentially, large drift over time. Due to the scanning method, horizontal displacements may occur. Usually, the images are not represented in perceptual color spaces, restricting the use of classical photometric invariances, such as in [10], but often there exist multiple signal and structural channels. On top of that, due to technical limitations, high speed imaging can often only be realized on narrow field of views (FOV).

Our image registration method builds on those observations: Due to the absence of different motion layers as well as of large displacements, we do neither need strategies for large displacements as in [6] nor a regularizer that preserves discontinuities as in [8]. Also, the assumptions of elastic regularizers which penalize the divergence of the OF field, e.g. compare [11], are violated by recordings during drug injection, see Figure 1 (F), where the displacement field has a high divergence after injection. Therefore, we quadratically penalize deviations from smooth displacements fields. In our previous work [12], we developed a motion compensation strategy for 1D linescans (e.g. [13]). We extend this approach to compensate recordings with narrow FOV and implement non-uniform warping in the optimization where we perform more warping steps along the larger image dimension.

Low-pass filtering is an important pre-processing step for local and global methods [14]. It makes the images derivable and integrates over temporal changes, while removing image noise. We found convolution with a 3D Gaussian kernel together with subquadratic penalization, to be robust enough to deal with the noise in the benchmark data, e.g., compare the layer 5 data (see supplemental Figure 3). Previous work has shown, that registration on structural channels does not necessarily result in much better registration performance [5]. However, in the variational framework, we can minimize the joint energy of the structural and the functional channels.

This has the advantage that, in theory, we get a better SNR, if the same structures are visible in both channels and, additionally, considering the aperture proplem, we have potentially more information on the motion, if disjunct structures are visible in

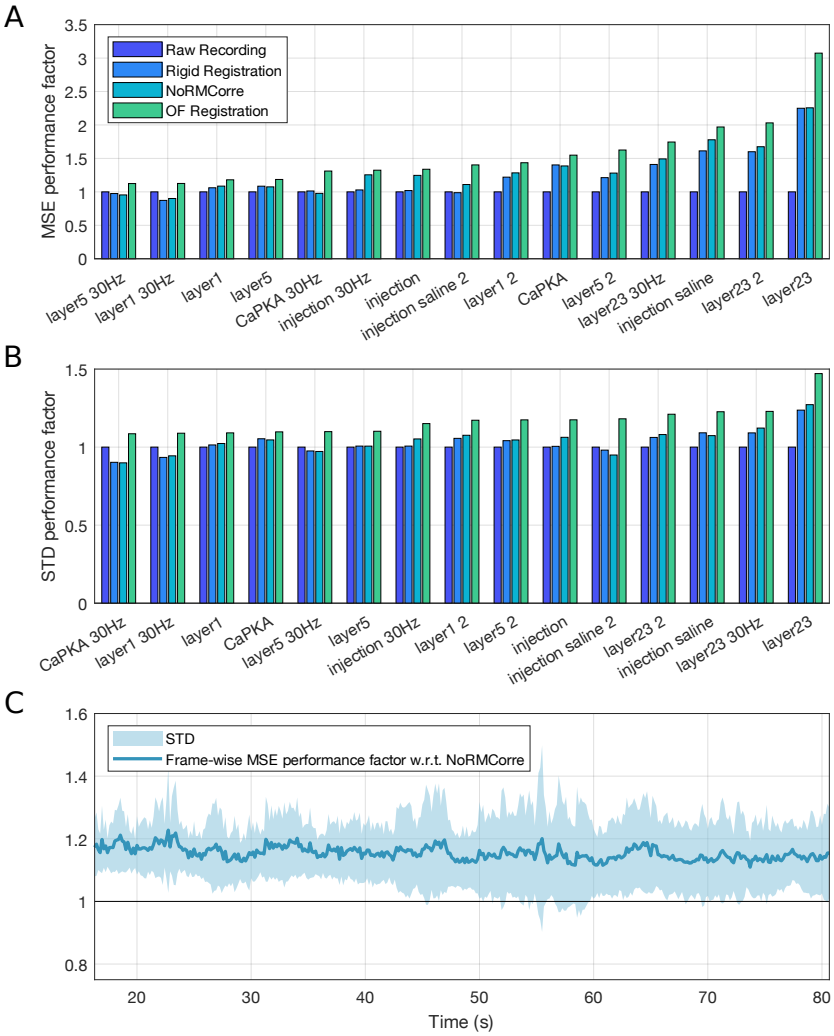

Figure 2: Application of Flow-Registration to 15 different datasets. We use 2-Photon recordings from layer 5, layer 2/3, layer 1 and three sequences during in vivo drug or saline injection at 6.2 Hz and 30.9 Hz. In all applied metrics, Flow-Registration performs consistently better than rigid registration and NoRMCorre (see Supplemental material for *peak signal to noise ratios* (PSNR)). The performance measures are averaged ratios of the raw recording and the compensated recordings with *mean squared error* (MSE) (A) and temporal *standard deviation* (temporal STD) (B). The performance factor indicates how much higher the MSE or temporal STD is in the raw recording, so higher values correspond to better compensation. (C) contains the frame-wise ratio of NoRMCorre and Flow-registration MSE of all 6.2 Hz datasets with 500 frames (layer1 - layer 5, injection, injection saline 2). MSE has been computed with respect to the first frames of each recording (see Online Methods for details). The frame-wise performance is significantly better for Flow-Registration ($p < 0.00001$, paired, two-sided Wilcoxon signed rank test). The performance on the datasets has been sorted with respect to the performance of Flow-Registration.

the different image channels, such as in the data in supplemental Figure 3 (A)-(D). Additionally, the energy functional can naturally incorporate a manual weighting term, such as ROIs, to enforce a low value of the dataterm around important image structures. For a fast approximation of the solution, the minimum warping depth can be used to

define the maximum resolution at which the solution is computed. We found around 10x speedup to be possible with similar compensation quality (avg PSNR -53.90 with the precise and -53.91 with the approximated solution, min level = 6, compare supplemental Video 1).

We can report state-of-the-art registration results in terms of reduced temporal standard deviation and MSE ratios, when comparing our method with NoRMCorre [5] and rigid image registration, see Figure 2. NoRMCorre has been adopted in recent toolboxes for 2P and calcium imaging such as CaImAn (2019) [15], EZcalcium (2020) [16] or Begonia (2021) [17] and can be considered the state of the art for the alignment of 2-photon image sequences.

In terms of computation time, the MATLAB toolbox is similar to existing methods and even faster with approximated solutions. On a single channel version of the 500 frame injection sequence (consumer workstation, 12 cores @ 3.8 GHz, 64GB of memory), our method (50 iterations, no pre-processing / IO, min level 0) takes 115 s and only 15 s with min level 6, while NoRMCorre (grid size 32, one iteration) takes 160 s and with grid size 16 even 763 s (setting used for the benchmark solution). Deep Learning based approaches could unlock real time application of our method, but even state-of-the-art self-supervised methods often require annotated data as initialization [18]. Flow-Registration has already been applied to many state-of-the-art 2-photon imaging recordings. The explicit, high-accuracy estimation of displacements can be used to generate datasets for the training of efficient, deep learning based motion compensation methods in the future.

Generally, a drawback of 2D motion compensation approaches for 2-photon imaging is the lack of z-shift correction. While there exist methods for high-speed, online 3D compensation [19], they require complicated setups and the current generation is limited to rigid 3D motion compensation — which might therefore benefit from refinement with a method such as Flow-Registration.

The solutions presented in this paper solve the pre-processing problem of motion contamination in microscopy and multichannel video recordings. Our core software design paradigm was the easy yet versatile integration into different workflows and toolboxes for 2-Photon imaging. We have developed a MATLAB toolbox that supports common file formats such as MDF, tiff image stacks, MATLAB mat files or hierarchical dataformat (HDF5) files in single file or multichannel configurations. The image IO is designed in a modular, object oriented way, such that the toolbox can easily be extended with new data formats and embedded pre-processing. The code is memory efficient and allows to compensate bigdata recordings with specified pre-processing methods and on-the-fly binning. The toolbox allows a variable number of channels which makes it suitable for multi-spectral recordings. The computationally heavy code is written in C++ which potentially allows the implementation of Python wrappers in the future. The ImageJ / FIJI plugin builds on Imglib2 [20] library which supports most common image formats through the bio-formats plugin. The implementations support weighted, multichannel input with weight masks for example to enforce higher weight on the dataterm inside of ROIs. The ImageJ / FIJI plugin is integrated with the MATLAB toolbox, such that the plugin can export parameters and reference frame configurations, that can be loaded as MATLAB bulk registration jobs.

1 Methods

Methods and any associated references are available in the online version of the paper.

2 Acknowledgements

This work was partially funded by the Japan Society for the Promotion of Science (JSPS) with the Summer Program 2019 under grand number SP19305 and partially conducted at the *Optical Neuroimaging Unit* under Bernd Kuhn at the Okinawa Institute of Science and Technology Graduate University. The authors would like to thank Miles J. Desforges, Mohamed Eltabbal and Lars Haab for code testing and valuable feedback and discussions.

3 Author Contributions

Philipp Flotho: Software, Methodology, Conceptualization, Data Curation, Formal Analysis, Visualization, Writing-original draft; **Shinobu Nomura**: Data Curation, Investigation, Writing-review & editing, Validation; **Bernd Kuhn**: Conceptualization, Data Curation, Writing-review & editing, Supervision, Resources; **Daniel J. Strauss**: Conceptualization, Writing-review & editing, Supervision

4 Conflict of Interest

The authors declare that there is no conflict of interest.

References

[1] Helmchen, F. & Denk, W. Deep tissue two-photon microscopy. *Nat Methods* **2**, 932 (2005).

[2] Theer, P., Kuhn, B., Keusters, D. & Denk, W. Two-photon microscopy and imaging (2006).

[3] Horn, B. K. P. & Schunck, B. G. Determining optical flow. *Artif Intell* **17**, 185–203 (1981).

[4] Greenberg, D. S. & Kerr, J. N. D. Automated correction of fast motion artifacts for two-photon imaging of awake animals. *J Neurosci Methods* **176**, 1–15 (2009).

[5] Pnevmatikakis, E. A. & Giovannucci, A. Normcorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *J Neurosci Methods* **291**, 83–94 (2017).

[6] Chen, Z., Jin, H., Lin, Z., Cohen, S. & Wu, Y. Large displacement optical flow from nearest neighbor fields. In *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*, 2443–2450 (2013).

[7] Baker, S. *et al.* A database and evaluation methodology for optical flow. *Int J Comput Vis* **92**, 1–31 (2011).

[8] Sun, D., Roth, S. & Black, M. J. Secrets of optical flow estimation and their principles. In *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*, 2432–2439 (IEEE, 2010).

[9] Brox, T., Bruhn, A., Papenberg, N. & Weickert, J. High accuracy optical flow estimation based on a theory for warping. In *Comput Vis ECCV*, 25–36 (Springer, 2004).

[10] Zimmer, H. *et al.* Complementary optic flow. In *Energy Minimization Methods Comput Vis Pattern Recognit*, 207–220 (Springer, 2009).

[11] Hermosillo, G., Chef d'Hotel, C. & Faugeras, O. Variational methods for multimodal image matching. *Int J Comput Vis* **50**, 329–343 (2002).

[12] Flotho, P. *et al.* Fast variational alignment of non-flat 1d displacements for applications in neuroimaging. *J Neurosci Methods* **353**, 109076 (2021).

[13] Roome, C. J. & Kuhn, B. Dendritic coincidence detection in Purkinje neurons of awake mice. *Elife* **9**, e59619 (2020).

[14] Bruhn, A., Weickert, J. & Schnörr, C. Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. *Int J Comput Vis* **61**, 211–231 (2005).

[15] Giovannucci, A. *et al.* Caiman an open source tool for scalable calcium imaging data analysis. *Elife* **8**, e38173 (2019).

[16] Cantu, D. A. *et al.* Ezcalcium: open-source toolbox for analysis of calcium imaging data. *Front Neural Circuits* **14**, 25 (2020).

[17] Bjørnstad, D. M. *et al.* Begonia—a two-photon imaging analysis pipeline for astrocytic ca2+ signals. *Front Cell Neurosci* **15**, 176 (2021).

[18] Liu, P., Lyu, M., King, I. & Xu, J. Selflow: Self-supervised learning of optical flow. In *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*, 4571–4580 (2019).

[19] Griffiths, V. A. *et al.* Real-time 3d movement correction for two-photon imaging in behaving animals. *Nat Methods* **17**, 741–748 (2020).

[20] Pietzsch, T., Preibisch, S., Tomančák, P. & Saalfeld, S. Imglib2 – generic image processing in java. *Bioinformatics* **28**, 3009–3011 (2012).

5 Online Methods

5.1 Animals

Experiments were approved by the OIST Institutional Animal Care and Use Committee, and performed in and AAALAC International accredited facility. Animals were maintained in a 12 h / 12 h light/dark cycle at 22 °C, with food and water available *ad libitum*.

5.2 Recombinant viruses

The adeno-associated virus (AAV) encoding the GAkdYmut PKA activity sensor [21] was custom made and produced by the vector core facility of Pennsylvania University (AAV2/1-hSyn-GAkdYmut-hGH, titer: 4×1014 gc/ml), and mixed with AAV2/1-hSyn-TurboRFP-WPRE encoding the red fluorophore RFP (titer: 4×1013 gc/ml, same supplier) at a ratio 1 : 1.

5.3 Expression of GAkdYmut and TurboRFP in cortex

Viral transfer of the indicator gene into cortical neurons of the mouse was performed as described before (Nomura et al. [22]). C57/BL6 mice (2-month-old) were anesthetized with a mixture of medetomidine (0.3 mg/kg), midazolam (4 mg/kg) and butorphanol (5 mg/kg). After performing a craniotomy over somatosensory cortex (AP −1.5 mm, ML 1.7 mm, DV −0.6-0.7 mm from bregma), 70-140 nl of a 1 : 1 mixture of AAV2/1-hSyn-GAkdYmut-hGH and AAV2/1-hSyn-TurboRFP-WPRE was injected in layer V at a rate of 10 nl/min. A chronic cranial window with a silicon access port (5 mm glass coverslip) was mounted as described Roome and Kuhn [23], [24]. At the end of the surgery, mice received atipamezole (0.3 mg/kg) for recovery from anesthesia, and buprenorphine (0.1 mg/kg) for pain relieve. Five to eight weeks after the AAV injection, mice were head-fixed for imaging experiments performed under anesthesia with 1% isoflurane or awake.

5.4 In vivo imaging in cortex

A combined wide-field / two-photon microscope (MOM, Sutter Instruments) with a femtosecond-pulsed Ti:sapphire laser (Vision II, Coherent) was used. To increase the point spread function of excitation the back aperture of the 25× water immersion objective (Olympus) was underfilled (spatial resolution 1 µm × 1 µm × 4 µm). The collar of the objective was adjusted to correct for the window glass thickness (170 µm). Simultaneous excitation of GAkdYmut (GFP-based single fluorophore sensor, Bonnot et al. [21]) and TurboRFP was performed at 950 nm with a typical power of 5-11 mW. Fluorescence was detected in two channels by GaAsP photomultipliers (Hamamatsu) in spectral windows 490−550 nm (GAkdYmut) and 600-700 nm (TurboRFP), separated by a 560 nm dichroic mirror (all Semrock). The microscope was controlled by a commercial software (MScan, Sutter Instruments). Sampling rate was 30.9 frame/s with 512 × 512 pixel, corresponding to a field of view of 375 µm × 375 µm. Saline (0.9% NaCl) with Alexa592 (1 µM) or with additional drug (propranolol 10 mM) was injected under anesthesia. Pressure injection was performed through the silicon access port using a glass pipette beveled to a diameter of 5-10 µm opening.

5.5 Benchmark Datasets and Evaluation

We used multiple datasets recorded with the described setup. The dataset names indicate the imaging depth (see supplemental Figure 3), where deeper layers usually correspond with lower SNR. For the layer 1-5 recordings, timepoints during movement onset were selected, as those events usually correspond with movement artifacts in the recording. For the datasets during drug and saline injections, we selected the time points around the injection events (e.g. compare Figure 1, (E)). From each dataset, we selected 2500 frames (80.9 s). A common approach for the evaluation of such data is temporal binning to increase SNR. We evaluated our methods on the datasets with binning over five frames (6.2 Hz, 500 frames total), as well as on a subset of the raw recordings, indicated by the suffix 30 Hz. For the dataset *saline injection*, in total 5 frames from the beginning and end of the experiment were excluded due to artifacts. We used real-world, low SNR datasets for the evaluation of our method without ground truth displacements, therefore, metrics such as endpoint or angular error are not applicable. The same holds true for perception based metrics due to the overall small movements and high image noise in the data. For the evaluation, we used reference based PSNR (see supplemental Table 2), MSE as well as temporal STD (see Figure 2) on 2D Gaussian lowpass filtered ($\sigma = (3, 3)^\top$) versions of the data. As reference, we used the temporal average of the each recording over the same frames that have been used as reference for the motion compensation. We report MSE and STD performance factors which indicate how much higher the MSE or STD value is for the raw sequence (see Figure 2 (A) and (B)) or how much higher the MSE is for NoRM-Corre when compared to Flow-Registration on a frame-wise basis (see Figure 2 (C)). We excluded all frames that contributed to the reference, as well as a boundary of 25 pixels. The boundary is motivated by the large maximum displacements encountered in the injection sequences. Lowpass filtering reduced the influence of image noise on the results and the boundary assures that only valid image regions are considered in the evaluations. For the compensation of all sequences, the first 100 frames (500 frames 30.9 Hz) were used as reference and supplied to the respective method.

5.6 OF Method

The applied motion estimation method is a robust variational OF method with gradient constancy, robust, separate channel penalization of the dataterm and first-order, isotropic, flow-driven regularizer. We normalize the dataterm according to Zimmer et al. [10] and apply the recommended good practices proposed by Sun et al. [8]. The regularizer and dataterm are penalized with a generalized Charbonnier penalty function which for $\epsilon > 0$ is given by $\Psi_a(x) = (x^2 + \epsilon^2)^a$, for $a > 0$ [8]. For the regularizer, larger values of a reduce discontinuities in the flow field (staircasing artifacts) and for the dataterm, smaller values of a reduce the influence of outliers. For 2-Photon recordings, we set the smoothness a to 1, which results in quadratic penalization. Note, that for $a = 0.5$ we get a regularized ℓ^1 norm, while the function becomes non-convex for values of $a < 0.5$. However, Sun et al. encourage a choice of $a = 0.45$ on the Middlebury benchmark, which we apply for the dataterm here. For the numerical approximation, we follow the framework of Bruhn et al. [25], Brox et al. [9] and Papenberg et al. [26] and discretize the *Euler Lagrange equations* in the compact *motion tensor* notation to solve them with an iterative multiscale solver (downsampling factor $\eta \in (0, 1)$) with

lagged non-linearities (update every 5 iterations). We use bicubic interpolation for the warping steps and 5×5 median-filtering (mirror boundary) of the flow increments for each level to increase accuracy as suggested by Sun et al. [8].
The main parameters of our method are the regularization parameter $\alpha = (\alpha_1, \alpha_2)^\top$, the penalizer power a, the warping depth, the downsampling factor η and the kernel size $\sigma = (\sigma_1, \sigma_2, \sigma_3)$.[2]. The choice of α presents the compromise of correctly registering smaller structures that deviate from the global motion direction and a globally smooth solution. With different values for α_1 and α_2, the smoothness term becomes anisotropic. This might support the estimation of motion artifacts induced by horizontal scanning. We scale alpha on each level of the scale space such that at a level i we use $\alpha = \alpha \cdot \eta^{-i/2}$. In practice, this made the result more robust under lower choices of α while enabling the compensation of high frequent jitter. To allow higher computational speed, the finest level for the OF calculation can be specified in the options. The Flow-Registration plugin implements this in the *Registration quality* setting, where only the highest quality setting calculates the solution on all levels. In the MATLAB toolbox, the minimum level can be set explicitly besides the abstract quality setting. With settings $\eta = 0.8$ and minimum level of 6, we get almost tenfold faster computation time on the injection sequence with almost the same accuracy (PSNR -53.90 vs. -53.91). With those parameters, the highest resolution at which the displacements are estimated are given by $512 \cdot 0.8^6 = 134$.

5.7 Data processing pipeline

The ImageJ / FIJI plugin makes use of the ImageJ file formats and therefore can resort to all supported file types. The MATLAB toolbox contains modular file readers and writer classes that can be automatically instantiated or supplied as parameters to an options object that defines a registration job. The file IO is designed for multi-channel processing and supplies 4D matrices in the format height×width×channel×time independently of the actual data representation on disk. To compensate a recording, the file reader supplies batches of size n with on-the-fly binning to the Flow-Registration engine which are then concurrently compensated. The average displacement of the last frames is used to initialize the lowest pyramid level of the displacement estimations in the following batch.
Reference frames can either be supplied directly or are computed from a specified set of frames, where the set of frames is aligned with respect to the temporal average and then temporally averaged again. For the initial registration, α and σ are increased in size to make the result more robust under noise and reduce overfitting. The reference and data are normalized with respect to the 3D Gaussian filtered reference frames. Joint normalization is performed by default but channel-wise normalization is supported as well.
The displacements of each frame with respect to the reference on the lowpass filtered sequences are computed and the raw frames are then registered via backwards warping with bicubic interpolation, where out-of-bounds values are replaced with the values from the reference. To reduce quantization due to the interpolation, the results are stored with double precision as default but can also be stored with the precision of the input file.
Batch processing is possible via the batch processor, where a list of filenames is supplied that will be compensated either with individual references or the same reference, when

dealing with recordings with the same ROI or recorded object. In the current version, the ImageJ / FIJI plugin can only be used for the compensation of short sequences (for example the 500 frame 6.2 Hz datasets or 2500 frames 30.9 Hz e.g. with stride 10 and 50 resp. for fast parameter testing), the extension for virtual stacks is planned for the future. To run the plugin, it needs to be installed in ImageJ / FIJI via Plugins→Install Plugin which adds a Flow-Registration entry under Plugins. For the MATLAB toolbox, a C++ compiler is required and the code has been tested for MATLAB R2018a onwards. The folder *demos* in the MATLAB toolbox contains scripts to reproduce the video results and quantitative results presented here as well as examples on how to use the code for different application scenarios. The *jupiter* demo compensates an amateur jupiter recording and demonstrates different aspects of multi-channel tiff and ROI processing as well as a potential application beyond the scope of neuroimaging.

5.8 Parameter Selection and Quality Metrics

For Flow-Registration, parameters were chosen based on visual inspection. For NoRMCorre, parameters from the demo scripts where modified based on the statistics calculated from the OF results, such as maximum displacements. NoRMCorre incorporates different regularization approaches that makes it more robust under noise and implicitly define the properties of the compensated displacements. There are a total of 7 parameters for regularization and 2 for subpixel refinement against the single regularization parameter α and additional solver specific parameters for Flow-Registration, while subpixel accuracy is natively supported. We found it necessary to modify 4 of the NoRMCorre parameters for our datasets (see Table 1). Dynamic update of the reference was turned off due to the reference based metrics that were used for evaluation. In practice, on the injection sequences, this reduced drift, but slightly increased jitter at the same time.
Flow-Registration was run for all datasets with default parameters and $\alpha = 1.5$, for the 6.2 Hz datasets with $\sigma = (1, 1, 0.1)^\top$ and on the 30.9 Hz datasets with $\sigma = (1, 1, 0.5)^\top$. For the injection sequence, additionally the channel weight was set to $(1.15, 0.85)^\top$. The batch size was set to the size of the datasets.

5.9 False Color Representation

The microscopy images presented in this paper aim to visualize artifacts in the average frames caused by residual motion in two imaging channels. Therefore, we make use of an inclusive (linear) false color representation that allows distinction of the channels both with deuteranopia and protanopia. Given the minimum intensity values $l = (l_1, l_2)^\top$ and maximum intensity values $h = (h_1, h_2)^\top$ from the temporal average of the raw recordings, the color $(R, G, B)^\top$ at a given image position $f(x, y) = (p_1, p_2)^2$ is calculated as

$$\begin{aligned} R &= (p_2 - l_2)/(h_2 - l_2) \\ G &= 0.5 \cdot (p - l)^\top \cdot \left((h_1 - l_1)^{-1}, (h_2 - l_2)^{-1}\right)^\top \\ B &= (p_1 - l_1)/(h_1 - l_1) \end{aligned}$$

This means that the reference normalized red imaging channel is mapped onto B, the green channel onto R and G is the average of the other two. Therefore, turquoise areas

dataset	grid_size	max_shift	max_dev	update_template
layer 1	$(32, 32)^\top$	–	–	false
layer 1 30Hz	$(64, 64)^\top$	–	–	false
layer23	$(32, 32)^\top$	–	–	false
layer23 30Hz	$(64, 64)^\top$	–	–	false
layer5	$(64, 64)^\top$	–	–	false
layer5 30Hz	$(128, 128)^\top$	–	–	false
CaPKA	$(32, 32)^\top$	–	–	false
injection	$(16, 16)^\top$	25	$(25, 25)^\top$	false
injection	$(16, 16)^\top$	25	$(25, 25)^\top$	false
injection 30Hz	$(64, 64)^\top$	25	$(25, 25)^\top$	false
injection saline	$(128, 128)^\top$	10	$(26, 26)^\top$	false
injection saline 2	$(32, 32)^\top$	15	$(25, 25)^\top$	false

Table 1: NoRMCorre parameters used for the compensation of the benchmark data. All unmentioned parameters have been kept in the default setting of the MATLAB code (June 2021). If the second version of the dataset is not mentioned, parameters match with the first version.

correspond channel 2 of the input, orange areas to channel 1 and white areas indicate that both channels are active.

5.10 Code and Data Availability

Documentation, the MATLAB code for Flow-Registration and the ImageJ Plugin can be found on the GitHub repository `https://github.com/phflot/flow_registration`. The version used for the evaluations in this paper is supplied as supplemental code. It contains MATLAB scripts to perform motion compensation with the parameters reported here as well as the precompiled ImageJ / FIJI plugin. The complete benchmark dataset used in this work is available as *2-Photon Imaging Motion Dataset* on Dryad.

Online Methods References

[21] Bonnot, A. *et al.* Single-fluorophore biosensors based on conformation-sensitive gfp variants. *FASEB J* **28**, 1375–1385 (2014).

[22] Nomura, S. *et al.* Combined optogenetic approaches reveal quantitative dynamics of endogenous noradrenergic transmission in the brain. *iScience* **23**, 101710 (2020).

[23] Roome, C. J. & Kuhn, B. Chronic cranial window with access port for repeated cellular manipulations, drug application, and electrophysiology. *Front Cell Neurosci* **8**, 379 (2014).

[24] Roome, C. J. & Kuhn, B. Voltage imaging with annine dyes and two-photon microscopy. In *Multiphoton Microscopy*, 297–334 (Springer, 2019).

[25] Bruhn, A. *Variational optic flow computation: accurate modelling and efficient numerics*. Ph.D. thesis, Saarland University (2006).

[26] Papenberg, N., Bruhn, A., Brox, T., Didas, S. & Weickert, J. Highly accurate optic flow computation with theoretically justified warping. *Int J Comput Vis* **67**, 141–158 (2006).

6 Supplemental Figures

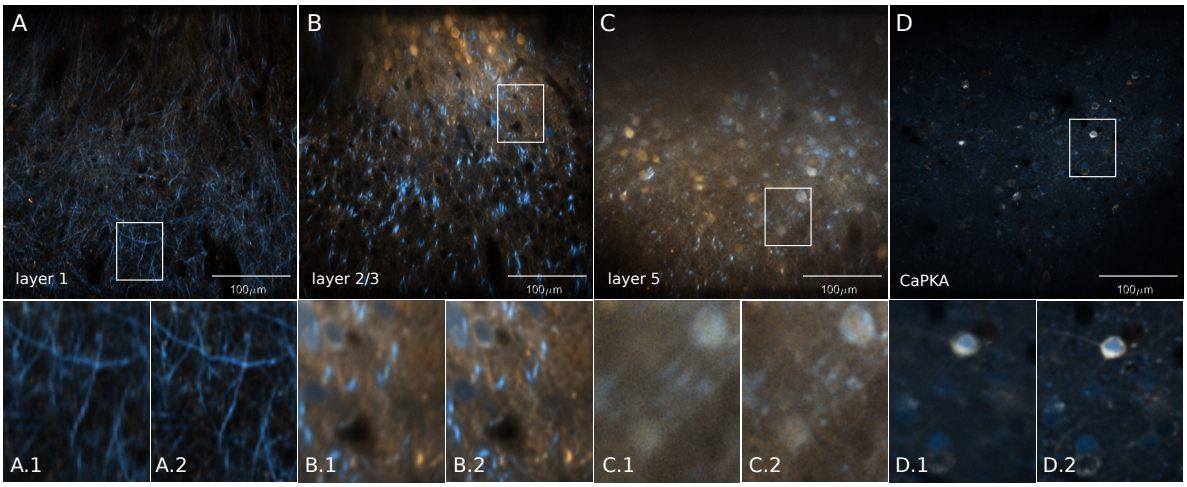

Figure 3: The datasets used for the evaluation besides the injection sequences. Temporal average of raw recordings (A.1-D.1) and after application of Flow-Registration (A.2-D.2) during motion.

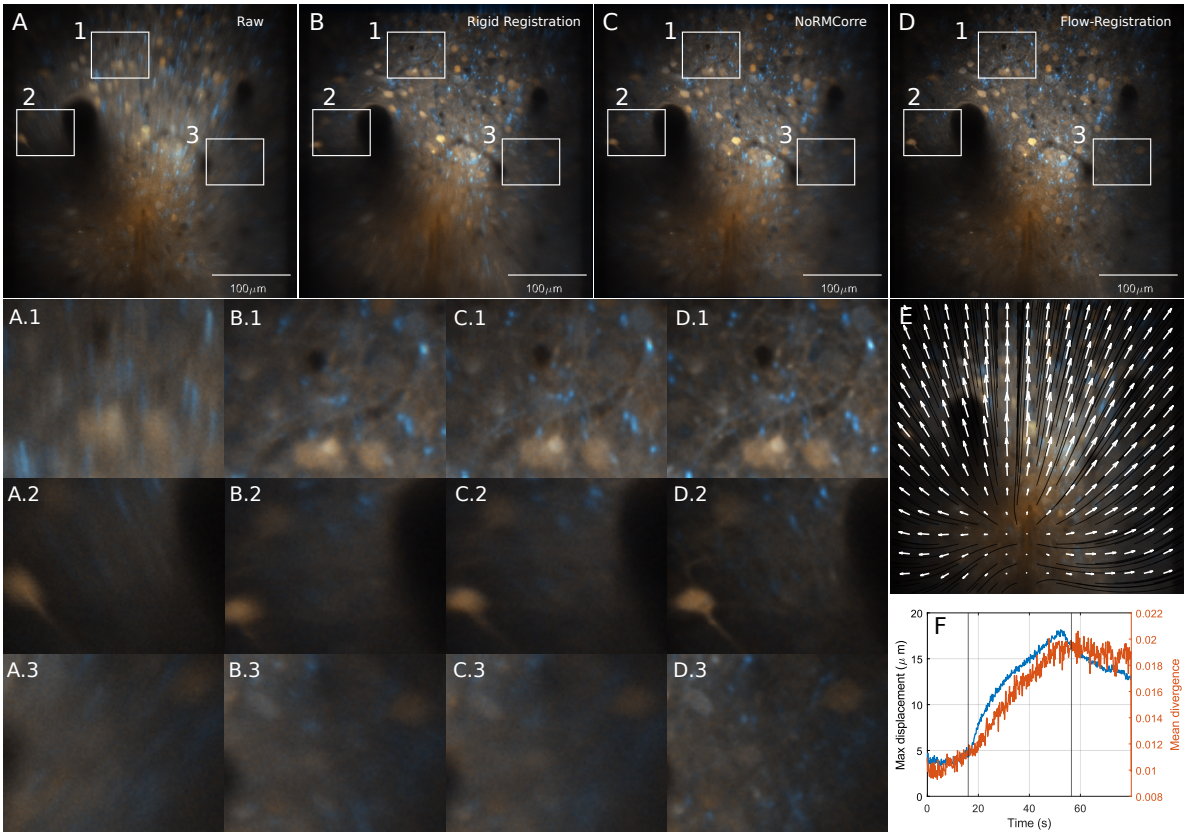

Figure 4: Qualitative results of the first in vivo saline injection sequence (dataset *injection saline*). Average of (A) raw images, (B) after rigid registration, (C) after registration with NoRMCorre, and (D) after Flow-Registration. The tissue expands from the injection point (E) resulting in large displacements as well as in a high divergence (F) in the displacement field. Flow-Registration can recover fine structures with much more detail, allowing region-of-interest selection along dendritic structures (blue), while the blur in (A-C) indicates residual motion.

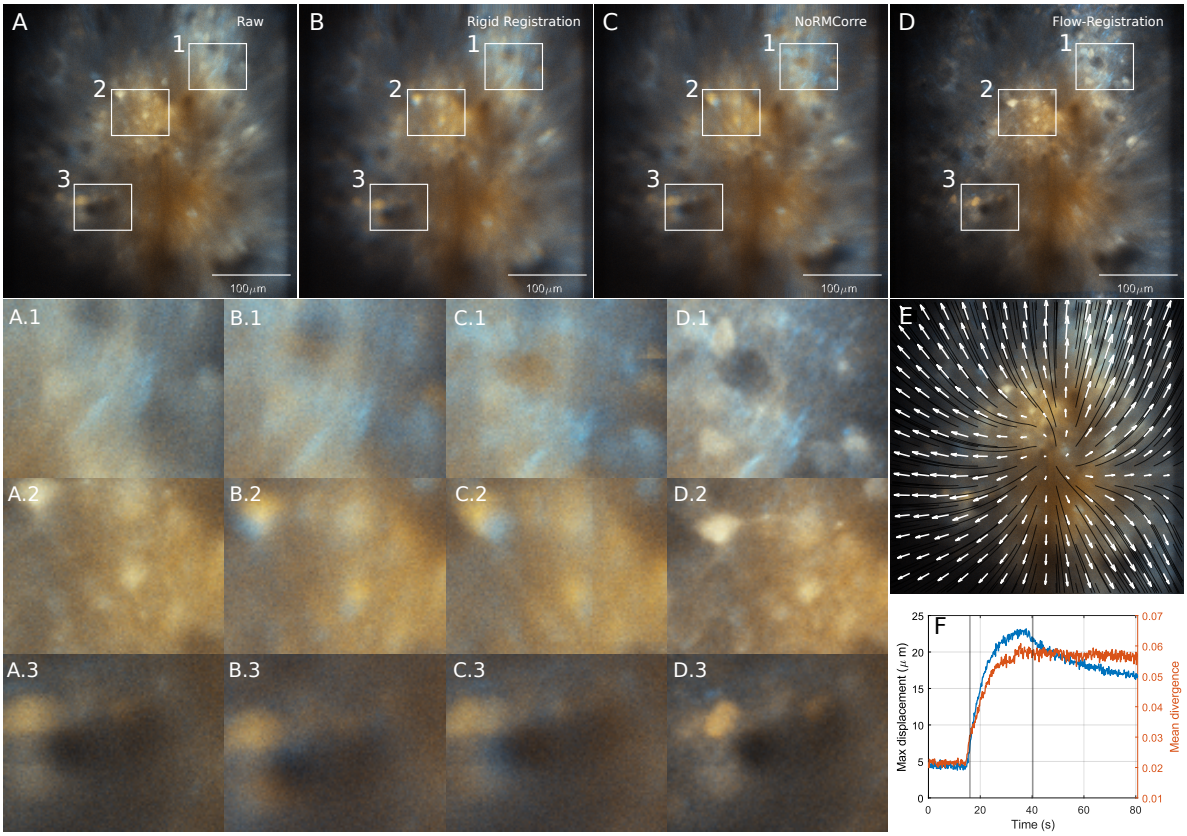

Figure 5: Qualitative results of the second in vivo saline injection sequence (dataset *injection saline 2*). Average of (A) raw images, (B) after rigid registration, (C) after registration with NoRMCorre, and (D) after Flow-Registration. The tissue expands from the injection point (E) resulting in large displacements as well as in a high divergence (F) in the displacement field. Flow-Registration can recover fine structures with much more detail, allowing region-of-interest selection along dendritic structures (blue), while the blur in (A-C) indicates residual motion.

	Raw	Rigid	NoRMCorre	Flow-Registration
CaPKA	-10.10	-8.81	-8.82	**-8.47**
CaPKA 30Hz	-14.22	-14.59	-14.64	**-13.32**
injection	-57.35	-56.73	-54.95	**-53.90**
injection 30Hz	-57.93	-57.42	-56.11	**-55.35**
injection saline	-18.11	-16.06	-15.53	**-15.01**
injection saline 2	-17.70	-17.69	-17.11	**-16.15**
layer1	-4.54	-4.37	-4.28	**-3.93**
layer1 2	-8.74	-8.19	-8.01	**-7.61**
layer1 30Hz	-11.17	-11.77	-11.65	**-10.67**
layer23	-19.19	-16.44	-16.40	**-15.01**
layer23 2	-7.01	-5.35	-5.11	**-4.27**
layer23 30Hz	-23.08	-21.78	-21.49	**-20.75**
layer5	-6.14	-5.83	-5.87	**-5.44**
layer5 2	-7.54	-6.95	-6.76	**-5.74**
layer5 30Hz	-12.55	-12.66	-12.75	**-12.05**

Table 2: Average PSNR values on each dataset for the different methods. Flow-Registration consistently outperforms the other methods. PSNR has been calculated with respect to the maximum value 2^{16}, experiment specific properties as well as the applied low-pass filtering (see section 5.5) as well as SNR of the raw data (e.g. 30.9 Hz vs 6.2 Hz) have an impact on the PSNR differences between the datasets. The best PSNR is put in bold.