



HHS Public Access

Author manuscript

Proc SPIE Int Soc Opt Eng. Author manuscript; available in PMC 2019 April 05.

Published in final edited form as:

Proc SPIE Int Soc Opt Eng. 2019 February ; 10949: . doi:10.1117/12.2512257.

Accurate registration of *in vivo* time-lapse images

Seyed M. M. Kahaki, Shih-Luen Wang, and Armen Stepanyants*

Department of Physics and Center for Interdisciplinary Research on Complex Systems,
Northeastern University, Boston, MA 02115, USA

Abstract

In vivo imaging experiments often require automated detection and tracking of changes in the specimen. These tasks can be hindered by variations in the position and orientation of the specimen relative to the microscope, as well as by linear and nonlinear tissue deformations. We propose a feature-based registration method, coupled with optimal transformations, designed to address these problems in 3D time-lapse microscopy images. Features are detected as local regions of maximum intensity in source and target image stacks, and their bipartite intensity dissimilarity matrix is used as an input to the Hungarian algorithm to establish initial correspondences. A random sampling refinement method is employed to eliminate outliers, and the resulting set of corresponding features is used to determine an optimal translation, rigid, affine, or B-spline transformation for the registration of the source and target images. Accuracy of the proposed algorithm was tested on fluorescently labeled axons imaged over a 68-day period with a two-photon laser scanning microscope. To that end, multiple axons in individual stacks of images were traced semi-manually and optimized in 3D, and the distances between the corresponding traces were measured before and after the registration. The results show that there is a progressive improvement in the registration accuracy with increasing complexity of the transformations. In particular, sub-micrometer accuracy (2–3 voxels) was achieved with the regularized affine and B-spline transformations.

Keywords

image registration; image alignment; non-rigid; time-lapse imaging; *in vivo* imaging; image stack

1. INTRODUCTION

Accurate registration is often required for the analysis and visualization of medical images. For instance, 3D optical microscopy imaging of a large region of interest is often performed by acquiring stacks of images that tile the region with overlaps [1, 2]. The same region, in addition, can be imaged multiple times in a time-lapse manner [3–5]. Because such imaging experiments are often done *in vivo*, registration is required to eliminate artifacts related to tissue translation, rotation, as well as linear and non-linear distortions. Three of the most common registration problems include (*i*) registration of image planes within individual stacks, (*ii*) registration of image stacks within a larger region of interest, and (*iii*) registration

* a.stepanyants@neu.edu; (617)373-2944; <http://www.northeastern.edu/neurogeometry/>.

of image stacks over time. More than a few algorithms have been developed over the years to solve these registration problems [6], however, in our experience, these algorithms do not reliably yield high accuracy required to address many biological questions. In particular, structural changes in the brain that accompany learning and memory formation can be visualized in *in vivo* imaging experiments, but automated detection and analyses of such changes are hindered by the relatively low accuracy of the existing registration methods. Because many plastic elements of synaptic connectivity, such as axonal boutons, dendritic spines, and terminal branches of neurons, have micrometer-scale dimensions, a sub-micrometer registration accuracy is required to disambiguate true structural changes in these elements from apparent changes brought up by tissue misalignment and distortion. In this study, we propose a feature-based registration method, capable of achieving such accuracy due in part to a two-step feature matching procedure. The method was validated on a dataset of two-photon laser-scanning microscopy images of mouse brain, which were acquired *in vivo* in 18 imaging sessions spanning the 68-day duration of the experiment [7].

2. METHOD

To register two time-lapse stacks of images, referred to as source and target, we detect intensity-based features in the two stacks, coarsely match the source and target features, refine these matches by eliminating outliers, and use the remaining matches to determine optimal registering transformations (Figure 1).

Features in this study are defined as small volumes ($9 \times 9 \times 9$ voxels) centered at local intensity maxima in the stacks (see Figure 1C). Although such features are not invariant under rotation or scaling, their use in this study is justified by the facts that the specimen is typically pre-aligned during image acquisition [7], and the tissue does not distort dramatically between imaging sessions (see e.g. Figure 1A). Other features, such as corners [8, 9], SIFT [10], and SURF [11], may be used in cases where the above assumptions are violated, so long as they provide dense and uniform coverage of the images. Prior to the detection of the maximum intensity voxels, the stacks are filtered with a Gaussian filter ($3 \times 3 \times 3$ voxels in size) to reduce the effects of noise. To achieve a reasonably uniform distribution of features in the images, we use a sequential algorithm in which after the detection of one local maximum, voxels belonging to the corresponding feature are eliminated from consideration, and the algorithm continues until there are no more features left. This process is applied to all image stacks in a given dataset, and the detected features, along with their positions, are saved. Large numbers of features are typically detected in every stack (~ 1200 features in the left image in Figure 1A), which, in theory, is beneficial for the registration accuracy. In practice, however, because the detected features are often very similar, conventional matching methods can lead to errors [12] and/or can be time-consuming if based on a combinatorial search strategy. To address these issues, we developed a two-step procedure in which features are first coarsely matched based on similarity, and then, matches are refined based on optimal transformations of randomly sampled subsets.

To determine coarse matches between m features in the source stack and n features in the target stack, we compute an $m \times n$ matrix of bipartite feature dissimilarities, S , in which an

element s_{ij} denotes the mean absolute difference between the mean normalized features i in the source stack and j in the target stack. In addition, we introduce a cost, $c > 0$, for leaving a feature without a match. Matrix S and cost c are then used as inputs to the Hungarian algorithm [13] to determine an optimal assignment for which the total dissimilarity of matched features plus the total cost of unmatched features is minimized:

$$\arg \min_A \left[\sum_{i=1}^m \sum_{j=1}^n s_{ij} a_{ij} + c \left(m + n - 2 \sum_{i=1}^m \sum_{j=1}^n a_{ij} \right) \right] \quad (1)$$

Binary adjacency matrix $A = (a_{ij})$ in this expression encodes feature matches, such that $a_{ij} = 1$ indicates that the feature i in the source stack is matched to the feature j in the target stack. Because feature assignment must be one-to-one, matrix A must contain no more than one non-zero element in every row and column.

The results of coarse feature matching typically contain erroneous matches referred to as outliers. For example, the pair of stacks shown in Figure 1A contains $N_H = 930$ coarse matches (lines), $\gamma = 0.2$ fraction of which are outliers (see e.g. Figure 1C). Cost c is the only parameter of the Hungarian algorithm. Its scale is defined by the values of feature dissimilarities, which are confined to the $[0, 1]$ range. The value of c can affect the fraction of outliers, but there is no need to precisely tune this parameter because outlier matches, which are inevitably present at this stage of the algorithm, will be eliminated in the feature refinement step. All results of this study were generated with $c = 0.5$.

Course matches could be used to find the optimal registering transformation, however, even a small fraction of outliers can lead to a significant reduction of the registration accuracy. Therefore, we eliminate outliers by using a random sampling procedure like the one described in [14]. In this procedure, a subset of k coarse matches is randomly sampled, and an optimal transformation is calculated to register this subset. The detected optimal transformation is then applied to all source features, and those that end up within d voxels from their corresponding target pairs are deemed to be inliers. This process continues until the maximum number of sampling steps, $N_{samples}$, is reached, and the largest set of inlier matches is used for the registration (Figure 1D).

In this study we consider and compare four types of optimal rigid and non-rigid transformations, $p' = T(p)$, where $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$:

$$\text{Translation: } p' = p + b \quad (2)$$

$$\text{Rigid: } p' = Rp + b; \quad R \in \text{SO}(3)$$

$$\text{Affine: } p' = Lp + b; \quad L \in \text{GL}(3)$$

$$\text{B-spline: } p' = p + \sum_{l,m,n=0}^{3,3,3} X \begin{bmatrix} p_x \\ n_x \end{bmatrix} + l - 1, \begin{bmatrix} p_y \\ n_y \end{bmatrix} + m - 1, \begin{bmatrix} p_z \\ n_z \end{bmatrix} + n - 1 B_l \left(\frac{p_x}{n_x} - \left\lfloor \frac{p_x}{n_x} \right\rfloor \right) B_m \left(\frac{p_y}{n_y} - \left\lfloor \frac{p_y}{n_y} \right\rfloor \right) B_n \left(\frac{p_z}{n_z} - \left\lfloor \frac{p_z}{n_z} \right\rfloor \right)$$

$$B_0(u) = \frac{1}{6}(1-u)^3; \quad B_1(u) = \frac{1}{6}(3u^3 - 6u^2 + 4); \quad B_2(u) = \frac{1}{6}(-3u^3 + 3u^2 + 3u + 1); \quad B_3(u) = \frac{1}{6}u^3$$

The B-spline transformation is implemented as described in [15]. Positive integers n_x , n_y , and n_z define the size of one B-spline cell (in voxels). The numbers of cells required to cover the entire stack in all dimensions are $\lceil N_x / n_x \rceil$, $\lceil N_y / n_y \rceil$, and $\lceil N_z / n_z \rceil$, where N_x , N_y , and N_z denote the stack dimensions in voxels, and brackets indicate the round-up operation. B-spline is governed by $3(\lceil N_x / n_x \rceil + 3)(\lceil N_y / n_y \rceil + 3)(\lceil N_z / n_z \rceil + 3)$ transformation parameters, which are included in the set of vectors $X \begin{bmatrix} p_x \\ n_x \end{bmatrix} + l - 1, \begin{bmatrix} p_y \\ n_y \end{bmatrix} + m - 1, \begin{bmatrix} p_z \\ n_z \end{bmatrix} + n - 1$.

Optimal transformations are determined by minimizing the mean square distance between the transformed source points $\{T(p_i)\}$ and the corresponding target points, $\{q_i\}$. The affine and B-spline transformations, which are non-rigid, are regularized with the Frobenius norm squared of the difference between the deformation gradient tensor, ∇T^T , and the identity matrix, I , averaged over the stack volume:

$$\arg \min_T \left(\frac{1}{k} \sum_{i=1}^k \|T(p_i) - q_i\|_2^2 + \frac{\mu}{N_x N_y N_z} \int \|\nabla T^T(p) - I\|_F^2 d^3 p \right) \quad (3)$$

In this expression, μ is referred to as the regularization strength. We use $\mu = 0$ for the translation and rigid transformations, and $\mu > 0$ for the affine and B-spline transformations. The optimization problem of Eq. (3) is solved analytically to find the optimal translation,

rigid, and affine transformations, and numerically with the gradient descent method in the case of B-spline.

The described feature refinement procedure is governed by parameters k , d , and $N_{samples}$. One pair of features is used for translation ($k = 1$), $k = 3$ is used for the rigid transformation, $k = 4$ for affine, and $k = 6$ for B-spline. Larger values of d generally lead to higher numbers of inlier matches, however, to avoid errors, d has to be less than the typical spacing between features in the stack. The value of $d = 2$ was used to produce the results of this study.

$N_{samples}$ must be large enough to ensure that at least one subset of sampled matches contains no outliers. This parameter can be estimated by using the number of coarse matches, N_H , resulting from the Hungarian algorithm and the fraction of outliers, γ . For this, we first calculate the probability of having no outliers in a single randomly chosen subset of k matches:

$$P = \binom{N_H(1-\gamma)}{k} / \binom{N_H}{k} \approx (1-\gamma)^k \quad (4)$$

The approximation made in this expression is valid for $N_H(1-\gamma) \gg k$, which is the case for the data considered in this study. By setting the probability of not finding a single all-inlier subset among $N_{samples}$ sampled subsets to 10^{-9} , i.e. $(1-P)^{N_{samples}} = 10^{-9}$, we obtain:

$$N_{samples} = \left\lceil \frac{-9}{\log(1 - (1-\gamma)^k)} \right\rceil \quad (5)$$

We note that Eq. (5) depends only on the fraction of outliers, γ , and the complexity of the registering transformation, k . Because the former cannot be estimated reliably by simply examining the results of course matching, we used $\gamma = 0.2$ for the data in Figure 1 and multiplied the output of Eq. (5) by a factor of 10. This procedure led to 130 samples for the translation, 290 for rigid, 400 for affine, and 690 for B-spline transformations.

Registration algorithm described in this study was implemented in MATLAB, and the source code is available at <https://github.com/neurogeometry/registrar>.

3. RESULTS

To evaluate the accuracy of the proposed registration algorithm, we applied it to *in vivo* time-lapse images of fluorescently labeled axons from mouse barrel cortex obtained as a part of another study [7]. The image stacks, $270 \times 270 \times 250 \mu\text{m}^3$, were acquired at a voxel volume of $0.26 \times 0.26 \times 0.80 \mu\text{m}^3$ with a 4-day interval in 18 imaging sessions. A subset of axons contained in the images was manually traced and optimized with the NCTracer software [16, 17], to be used for validation. Figures 2A, B show maximum-intensity projections of two image stacks acquired four days apart in the same brain region. A custom-built system was used to pre-align the animal's head during imaging [7], but residual misalignment is clearly

visible in the red-green superimposed images (Figure 2C). A similar trend is observed in the overlay of traces of the same axons acquired in different imaging sessions (see e.g. Figure 2D). This misalignment makes it difficult to detect and analyze changes in the same structures across imaging sessions.

Registration can significantly improve upon the initial pre-alignment of image stacks. This is illustrated in Figures 3A, B, which show the overlays of image stacks and traces from Figures 2C, D following the registration. To evaluate the registration accuracy, the same set of 22 axons was traced in 18 imaging sessions, and the average distances between the corresponding traces were calculated before and after the registration. The results show that significant improvement over the original alignment is achieved with all four registering transformation types (Figure 3C). The average distance between traces monotonically decreases with increasing complexity of the transformations, and the highest accuracy (2.3 voxels or 0.6 μm) was attained with the regularized B-spline transformation. We note that the optimal affine and B-spline transformations depend on the value of the regularization parameter, μ [see Eq. (3)]. However, this dependence is not very strong (Figure 3D), eliminating the need for precise tuning of this parameter. The value of $\mu = 1000$ was used to produce the results of Figure 3.

4. CONCLUSION

We present accurate rigid and non-rigid methods for registering 3D optical microscopy stacks of images acquired in a time-lapse manner. The registration accuracy gradually improved with increasing complexity of the transformations, and sub-micrometer accuracy was achieved with the regularized B-spline (Figure 3). The proposed method is not specific to time-lapse images; it can be adapted to other essential medical imaging problems, including alignment of image planes within individual stacks and registration of overlapping stacks acquired in large-scale (e.g. whole brain) imaging experiments.

ACKNOWLEDGMENTS

This work was supported by the NIH grant R01 NS091421.

REFERENCES

- [1]. Schindelin J, Arganda-Carreras I, Frise E et al., "Fiji: an open-source platform for biological-image analysis," *Nat Methods*, 9(7), 676–82 (2012). [PubMed: 22743772]
- [2]. Economo MN, Clack NG, Lavis LD et al., "A platform for brain-wide imaging and reconstruction of individual neurons," *Elife*, 5, e10566 (2016). [PubMed: 26796534]
- [3]. Holtmaat A, and Svoboda K, "Experience-dependent structural synaptic plasticity in the mammalian brain," *Nat Rev Neurosci*, 10(9), 647–58 (2009). [PubMed: 19693029]
- [4]. Trachtenberg JT, Chen BE, Knott GW et al., "Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex," *Nature*, 420(6917), 788–794 (2002). [PubMed: 12490942]
- [5]. Holtmaat A, Bonhoeffer T, Chow DK et al., "Long-term, high-resolution imaging in the mouse neocortex through a chronic cranial window," *Nature Protocols*, 4(8), 1128–44 (2009). [PubMed: 19617885]
- [6]. Sotiras A, Davatzikos C, and Paragios N, "Deformable medical image registration: a survey," *IEEE Transactions on Medical Imaging*, 32(7), 1153–1190 (2013). [PubMed: 23739795]

- [7]. Gala R, Lebrecht D, Sahlender DA et al., "Computer assisted detection of axonal bouton structural plasticity in in vivo time-lapse images," *Elife*, 6, (2017).
- [8]. Harris C, and Stephens M, "A combined corner and edge detector," *Alvey vision conference*, 15(50), 10–5244 (1988).
- [9]. Kahaki SMM, Nordin MJ, and Ashtari AH, "Contour-based corner detection and classification by using mean projection transform," *Sensors*, 14(3), 4126–4143 (2014). [PubMed: 24590354]
- [10]. Lowe DG, "Object recognition from local scale-invariant features," *The proceedings of the seventh IEEE international conference on computer vision*, 2, 1150–1157 (1999).
- [11]. Bay H, Ess A, Tuytelaars T et al., "Speeded-up robust features (SURF)," *Computer vision and image understanding*, 110(3), 346–359 (2008).
- [12]. Kahaki SMM, Nordin MJ, Ashtari AH et al., "Deformation invariant image matching based on dissimilarity of spatial features," *Neurocomputing*, 175, 1009–1018 (2016).
- [13]. Munkres J, "Algorithms for the assignment and transportation problems," *Journal of the society for industrial and applied mathematics*, 5(1), 32–38 (1957).
- [14]. Brown M, and Lowe DG, "Automatic panoramic image stitching using invariant features," *International journal of computer vision*, 74(1), 59–73 (2007).
- [15]. Rueckert D, Sonoda LI, Hayes C et al., "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Transactions on Medical Imaging*, 18(8), 712–721 (1999). [PubMed: 10534053]
- [16]. Gala R, Chapeton J, Jitesh J et al., "Active learning of neuron morphology for accurate automated tracing of neurites," *Frontiers in Neuroanatomy*, 8, (2014).
- [17]. Chothani P, Mehta V, and Stepanyants A, "Automated tracing of neurites from light microscopy stacks of images," *Neuroinformatics*, 9(2–3), 263–78 (2011). [PubMed: 21562803]

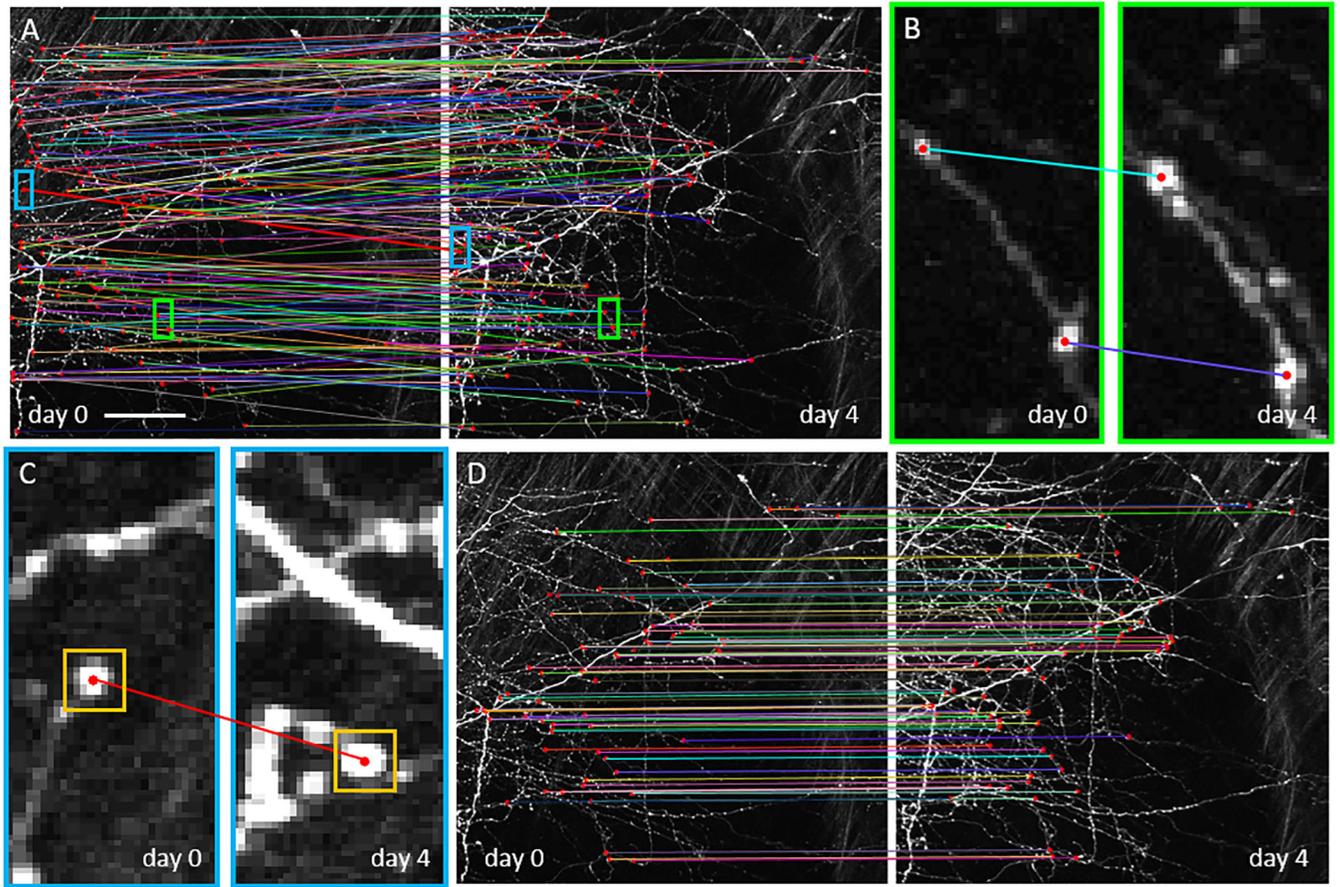


Figure 1.

Two-step feature matching method for accurate registration. A. Maximum intensity projections of two 3D stacks of images acquired *in vivo* with a 4-day interval in the same area of mouse barrel cortex (left and right images). Red dots denote the positions of detected feature centers. Only 15% of features are shown to avoid clutter. Coarsely matched features are indicated with lines. Matches at this step may contain outliers. B. Zoomed-in view of correctly matched features from (A). C. Same for incorrectly matched features (outliers). Yellow squares outline the feature boundaries. D. Outliers are eliminated with feature refinement. Scale bar in (A) corresponds to 50 μm for (A, D) and 4.3 μm for (B, C).

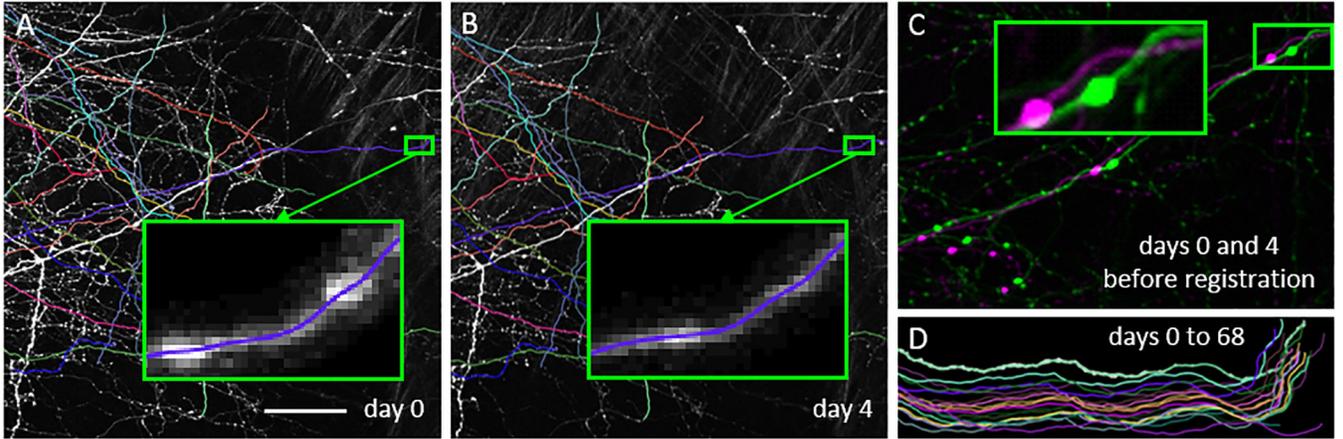


Figure 2.

Registration is often required for automated analyses of *in vivo* time-lapse images of neurons. A, B. Maximum intensity projections of two image stacks acquired with a 4-day interval showing fluorescently labeled axons of cortical neurons. Traces of some of the axons obtained with the NCTracer software are shown with colored lines. Insets in (A) and (B) show zoomed-in views of the same axon segment with some structural changes visible (bright varicosities in A). C. Superimposed stacks from (A) and (B), based on the pre-alignment done during the experiment, illustrate displacement and some distortion of the tissue between imaging sessions (red-green). D. Likewise, an overlay of traces of an axon segment imaged in 18 sessions shows a significant misalignment. Scale bar in (A) corresponds to 50 μm for (A, B), and 28 μm for (C, D).

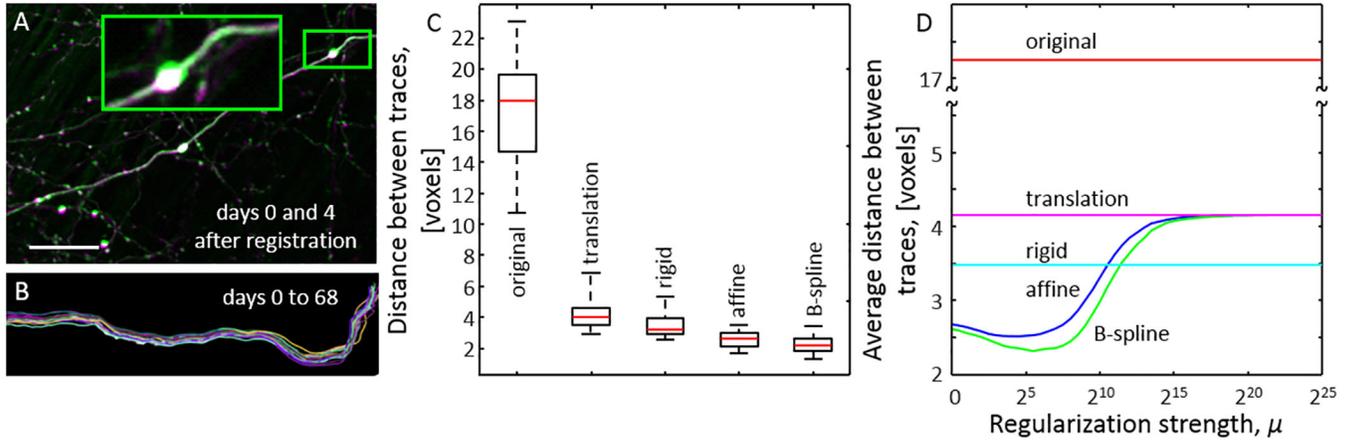


Figure 3. Validation of the registration procedure. A. After registration (B-spline), the red-green superimposed images from Figure 2C show a significant improvement in alignment. B. The same trend is observed for the axon traces from Figure 2D. Scale bar in (A) is 28 μm for (A, B). C. Box plots show the distances between traces of the same axons in subsequent time-lapse images before (original) and after registration (translation, rigid, affine, and B-spline). D. Average distances between time-lapse traces for the affine and B-spline transformations as functions of the regularization strength, μ . Corresponding distances based on the original images and images registered with the translation and rigid transformations are shown for comparison.