# Model-Based Refinement of Nonlinear Registrations in 3D Histology Reconstruction

Juan Eugenio Iglesias[1(✉)], Marco Lorenzi[2], Sebastiano Ferraris[1], Loïc Peter[5], Marc Modat[1], Allison Stevens[3], Bruce Fischl[3,4], and Tom Vercauteren[5]

[1] Medical Physics and Biomedical Engineering,
University College London, London, UK
[2] Epione Team, INRIA Sophia Antipolis, Sophia Antipolis Cedex, France
[3] Martinos Center for Biomedical Imaging, Harvard Medical School
and Massachusetts General Hospital, Boston, USA
[4] MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, USA
[5] Wellcome EPSRC Centre for Interventional and Surgical Sciences (WEISS),
University College London, London, UK

**Abstract.** Recovering the 3D structure of a stack of histological sections (3D histology reconstruction) requires a linearly aligned reference volume in order to minimize $z$-shift and "banana effect". Reconstruction can then be achieved by computing 2D registrations between each section and its corresponding resampled slice in the volume. However, these registrations are often inaccurate due to their inter-modality nature and to the strongly nonlinear deformations introduced by histological processing. Here we introduce a probabilistic model of spatial deformations to efficiently refine these registrations, without the need to revisit the imaging data. Our method takes as input a set of nonlinear registrations between pairs of 2D images (within or across modalities), and uses Bayesian inference to estimate the most likely spanning tree of latent transformations that generated the measured deformations. Results on synthetic and real data show that our algorithm can effectively 3D reconstruct the histology while being robust to $z$-shift and banana effect. An implementation of the approach, which is compatible with a wide array of existing registration methods, is available at JEI's website: www.jeiglesias.com.

## 1 Introduction

Combining histology with mm-scale volumetric images finds multiple applications in areas such as atlas building (e.g., [1]) or modeling the relationship between the information at micro- and macro-scale (e.g., [2]). Combining the two modalities requires histology reconstruction, i.e., registration of 2D histological sections to volumes to recover the lost 3D structure. The role of the reference volume is critical in reconstruction; in its absence, one can only resort to pairwise

M. Modat and T. Vercauteren—Currently with the School of Biomedical Engineering and Imaging Science, King's College London.

registration of the histological sections [3], which leads to $z$-shift (accumulation of registration errors) and banana effect (straightening of curved structures). In the remainder we assume that a reference volume is always available.

Most reconstruction methods assume that the nonlinear deformations in the histological sections occur within plane. Then, reconstruction can be decoupled into estimating a 3D transformation (often linear) between the histological stack and the volume, and estimating a set of nonlinear 2D transformations between each section and the corresponding resampled plane in the registered volume. An intermediate modality is sometimes used to estimate the 3D transformation; e.g., blockface photographs can be stacked and linearly registered to the reference volume [4]. Otherwise, the two problems can be solved iteratively, i.e., optimizing a 3D transformation with the 2D deformations fixed, and vice versa [5].

If the 3D registration is fixed, any 2D, nonlinear, intermodality registration algorithm can be used to align the histological sections to the corresponding resampled slices in the reference volume. However, this baseline approach often yields jagged reconstructions, as this registration problem is difficult to solve due to two reasons: the large and slice-varying differences in intensity and contrast profiles between the modalities; and the extensive geometric distortion introduced by sectioning and mounting of the tissue – including folding and tears.

Smoother, more accurate reconstructions, can be obtained by considering neighboring sections when computing the 2D registrations, e.g., a section is deformed to match not only the corresponding resampled slice, but also the sections right above and below [1]. This approach unfortunately inherits the efficiency limitations of coordinate descent, and is thus computationally expensive: it requires many passes over each slice, and the use of a cost function with three channels. For instance, 20–25 reconstruction iterations were required on average in [1], hence running 60–70 times slower than the baseline approach.

A faster alternative is to first compute 2D nonlinear registrations between neighboring histological sections, and then use them regularize the 2D transformations between the histology and the reference volume. Since images are not revisited to update registrations, such approaches are computationally efficient. For example, poor linear registrations between neighboring sections are corrected by registration to other neighbors in [6]. Other approaches seek smoothness in the $z$ direction with low-pass filtering in linear [5,6] or nonlinear transformation spaces [7]. However, the optimality of such *ad hoc* approaches is unclear; e.g., early stopping is required in [7] as it converges to a solution with banana effect.

Here we present a probabilistic model in the space of 2D nonlinear spatial deformations, which accommodates measurements (2D registrations) between arbitrary pairs of sections and slices, neighboring or not, within or across modalities. The measurements are assumed to be noisy observations of compositions of latent transformations, which interconnect all the images in the two datasets through a spanning tree. We then use Bayesian inference to estimate the most likely latent transformations that generated the observations. Compared to previous works: 1. We explicitly optimize a principled objective function, which achieves smooth registrations while minimizing $z$-shift and banana effect;

2. Model parameters are estimated in inference (no parameter tuning required); and 3. Thanks to approximate composition of stationary velocity fields (SVF) [8], the latent transformations can be globally and efficiently optimized for fixed model parameters.

## 2   Methods

### 2.1   Probabilistic Framework

Let $\{H_n(\boldsymbol{x})\}_{n=1,\ldots,N}$ be a stack of $N$ histological sections, where $\boldsymbol{x} \in \Omega$ represents the pixel locations over a discrete, 2D image domain $\Omega$. Let $\{M_n(\boldsymbol{x})\}_{n=1,\ldots,N}$ be the corresponding slices of the reference volume, which we will henceforth assume to be an MRI scan. We further assume that the MRI and histological stack have been linearly aligned (e.g., with [5]). Let $\{\mathcal{T}_n\}_{n=1,\ldots,2N-1}$ be a set of $2N - 1$ latent, noise-free, nonlinear, diffeormorphic spatial transformations, which yield a spanning tree interconnecting all the images in the two datasets. Although our algorithm is independent of the choice of tree, we use for convenience the convention in Fig. 1a: $\mathcal{T}_n$, for $n \leq N$, registers histological section $n$ to MRI slice $n$ (i.e., maps coordinates from MRI space to histology); and $\mathcal{T}_{N+n}$ maps MRI slice $n$ to MRI slice $n + 1$, thereby modeling the banana effect and $z$-shift.

Now, let $\{\mathcal{R}_k\}_{k=1,\ldots,K}$ be a set of $K$ diffeomorphic transformations between pairs of images in the dataset (*any* pair, within or across modalities), estimated with an image registration method. $\mathcal{R}_k$ can be seen as a noisy version of a transformation equal to the composition of a subset of (possibly inverted) latent transformations of the spanning tree $\{\mathcal{T}_n\}$. In general, $K$ will be several times larger than $N$, and we can use Bayesian inference to estimate the latent transformations $\{\mathcal{T}_n\}$ that most likely generated the observed registrations $\{\mathcal{R}_k\}$.

Our choice of modeling nonlinear spatial transformation with diffeomorphisms is motivated by the Log-Euclidean framework, in which transformations are parameterized in the Lie group of SVFs [8]. Let $\{\boldsymbol{T}_n\}$ and $\{\boldsymbol{R}_k\}$ be the SVFs of the transformations, whose Lie exponentials are the corresponding diffeomorphisms $\{\mathcal{T}_n\}$ and $\{\mathcal{R}_k\}$, i.e., $\mathcal{T}_n = \exp[\boldsymbol{T}_n]$ and $\mathcal{R}_k = \exp[\boldsymbol{R}_k]$. Then, it follows (one-parameter subgroup property) that the inverse of a transformation is equivalent to its negation in the Log-space, i.e., $\mathcal{T}_n^{-1} = \exp[-\boldsymbol{T}_n]$. Moreover, the composition of diffeomorphisms parameterized by SVFs is given by $\boldsymbol{T}_n \oplus \boldsymbol{T}_{n'} = \log[\exp(\boldsymbol{T}_n) \circ \exp(\boldsymbol{T}_{n'})]$, whose analytical solution is provided by the Baker-Campbell-Hausdorff (BCH) series. By truncating the Lie bracket and considering only its first order terms, the BCH series can be approximated by: $\mathcal{T}_n \circ \mathcal{T}_{n'} \approx \exp[\boldsymbol{T}_n + \boldsymbol{T}_{n'}]$. While this approximation theoretically only holds for relatively small deformations, it is commonplace in state-of-the-art registration methods (e.g., [9]), and also enables us to globally optimize the objective function with respect to the transformations in inference – see Sect. 2.2 below.

To model the noise in the registrations, we choose an isotropic Gaussian model in the Log-space, which, despite its simplicity, works well in practice

(e.g., [8]). Henceforth, we will assume that $\boldsymbol{T}_n$ and $\boldsymbol{R}_k$ are shaped as $|\Omega| \times 1$ vectors. Then, the SVFs of the registrations are independent and distributed as:

$$\boldsymbol{R}_k \sim \mathcal{N}\left(\sum_{n=1}^{2N-1} w_{k,n}\boldsymbol{T}_n, \sigma_k^2\boldsymbol{I}\right),$$

where $\mathcal{N}$ is the Gaussian distribution; $\sigma_k^2$ is the variance of the $k^{th}$ measurement; $\boldsymbol{I}$ is the identity matrix; and $W := (w_{k,n})$, with $w_{k,n} \in \{-1, 0, 1\}$, is a matrix encoding the latent transformations that the registrations $\{\mathcal{R}_k\}$ traverse. Therefore, $\boldsymbol{R}_k$ is a Gaussian distribution centered on the concatenation of latent transformations corresponding to the measurement. More specifically, $w_{k,n} = 1$ if $\mathcal{T}_n$ is part of the path traversed by $\mathcal{R}_k$ with positive orientation, $w_{k,n} = -1$ if it is part of it with negative orientation, and $w_{k,n} = 0$ otherwise. Therefore, if a measurement estimates a transform from MRI slice $n'$ to MRI slice $n'' \geq n'$, then $w_{k,n} = 1$, for $n' + N \leq n < n'' + N$, and 0 otherwise. If the measurement is from MRI slice $n'$ to histological section $n''$, then $w_{k,n''} = 1$ needs to be added to $W$. And if the measurement is between histological sections $n'$ and $n''$, then we need to set $w_{k,n'} = -1$ in $W$, as well (see example in Fig. 1b).

The probabilistic framework is completed by a model for $\sigma_k^2$. A simple choice that we found to work well in practice is:

$$\sigma_k^2 = c_k\sigma_c^2 + d_k\sigma_d^2,$$

where $c_k = 1$ when transformation $k$ is across modalities (0 otherwise); $d_k$ is the number of sections or slices that $\mathcal{R}_k$ traverses; and $\sigma_c^2$ and $\sigma_d^2$ are the associated variance parameters, which will be estimated in inference – details below.



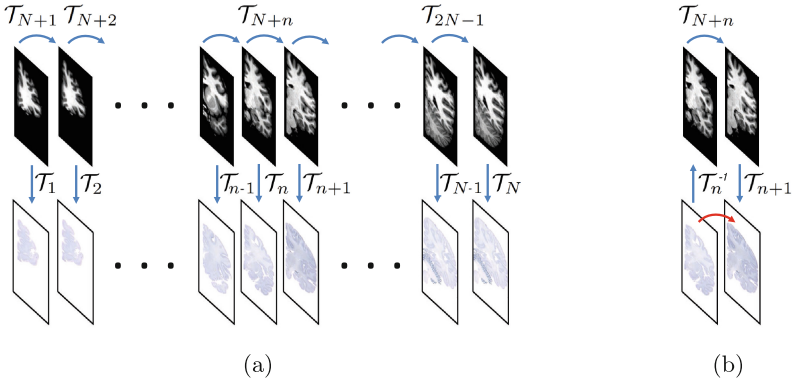(a)                                              (b)

**Fig. 1.** (a) Latent transformations $\{\mathcal{T}_n\}$ connecting the MRI and histological sections in the proposed model. A transformation between any pair of images can be written as the composition of a subset of (possibly inverted) transformations in $\{\mathcal{T}_n\}$. For example, the transformation represented by the red arrow in (b) can be written as: $\mathcal{T}_n^{-1} \circ \mathcal{T}_{N+n} \circ \mathcal{T}_{n+1}$ – or, in log-Euclidean framework, approximated as $\exp[-\boldsymbol{T}_n + \boldsymbol{T}_{N+n} + \boldsymbol{T}_{n+1}]$.

## 2.2   Inference: Proposed Method

In a fully Bayesian approach, one would optimize the posterior probability of the first $N$ transformations $\{\mathcal{T}_n\}_{n=1,\ldots,N}$ given the observations $\{\boldsymbol{R}_k\}$. However, computing such a posterior requires marginalizing over the remaining $N-1$ transformations and the variances $\sigma_c^2, \sigma_d^2$. A much simpler and faster inference algorithm can be derived by maximizing the joint posterior of all latent transforms $\{\mathcal{T}_n\}_{n=1,\ldots,2N-1}$ and variances $[\sigma_c^2, \sigma_d^2]^{\mathrm{T}}$, which is a good approximation because: 1. the uncertainty in the intramodality transformations ($n \leq N$) is much lower than that in the intermodality transformations ($n > N$); and 2. the two noise parameters can be accurately estimated, given residuals at $K|\Omega|$ locations. Using Bayes' rule and taking logarithm, we obtain the cost function:

$$\operatorname*{argmax}_{\{\boldsymbol{T}_n\},\sigma_c^2,\sigma_d^2} p(\{\boldsymbol{T}_n\}, \sigma_c^2, \sigma_d^2 | \{\boldsymbol{R}_k\}) = \operatorname*{argmax}_{\{\boldsymbol{T}_n\},\sigma_c^2,\sigma_d^2} p(\{\boldsymbol{R}_k\} | \{\boldsymbol{T}_n\}, \sigma_c^2, \sigma_d^2)$$

$$= \operatorname*{argmin}_{\{\boldsymbol{T}_n\},\sigma_c^2,\sigma_d^2} \sum_{k=1}^{K} \left[ |\Omega| \log\left[2\pi\left(c_k\sigma_c^2 + d_k\sigma_d^2\right)\right] + \frac{\|\boldsymbol{R}_k - \sum_{n=1}^{2N-1} w_{k,n}\boldsymbol{T}_n\|^2}{2\left(c_k\sigma_c^2 + d_k\sigma_d^2\right)} \right], \quad (1)$$

which we alternately minimize with respect to $\{\boldsymbol{T}_n\}$ and $[\sigma_c^2, \sigma_d^2]^{\mathrm{T}}$.

**Update for the Transformations:** Since we approximate composition by addition in the space of SVFs, optimizing Eq. 1 with respect to $\{\boldsymbol{T}_n\}$ is just a weighted least squares problem, with a closed form expression for its global optimum. Moreover, and since $W$ does not depend on pixel locations, the solution is given by a location-independent set of regression coefficients:

$$\boldsymbol{T}_n \leftarrow \sum_{k=1}^{K} z_{n,k} \boldsymbol{R}_k, \quad \text{with} \quad Z = \left(W^{\mathrm{T}} \operatorname{diag}(1/\sigma_k^2) W\right)^{-1} W^{\mathrm{T}} \operatorname{diag}(1/\sigma_k^2), \quad (2)$$

where $Z := (z_{n,k})$ is the matrix of regression coefficients. We note that *all* measurements can impact the estimation of every deformation.

**Update for the Variances:** We update the variances $\sigma_c^2$ and $\sigma_d^2$ simultaneously using a quasi-Newton method (L-BFGS). The gradient of the cost in Eq. 1 is:

$$\nabla_\sigma \mathcal{C} = \sum_{k=1}^{K} \left( \frac{|\Omega|}{c_k\sigma_c^2 + d_k\sigma_d^2} - \frac{E_k}{2(c_k\sigma_c^2 + d_k\sigma_d^2)^2} \right) [c_k, d_k]^{\mathrm{T}}, \quad (3)$$

where $E_k = \|\boldsymbol{R}_k - \sum_{n=1}^{2N-1} w_{k,n}\boldsymbol{T}_n\|^2$ is the energy of the $k^{th}$ residual.

**Practical Implementation:** We initialize the $\{\mathcal{T}_n\}$ with direct measurements, i.e., registrations $\{\mathcal{R}_k\}$ that map the same pairs of images as $\{\mathcal{T}_n\}$. Next, we compute the average squared error $S = (K|\Omega|)^{-1} \sum_k E_k$, and initialize $\sigma_c^2 = 3S$, $\sigma_d^2 = S/3$, such that $\sigma_c^2 \approx 10\sigma_d^2$. Finally, we iterate between updating $\{\mathcal{T}_n\}$ (with Eq. 2) and $[\sigma_c^2, \sigma_d^2]^{\mathrm{T}}$ (numerically with the gradient in Eq. 3). The algorithm often converges in 3–4 iterations (approximately an hour, including registrations).

# 3   Experiments and Results

## 3.1   Data

We validated our method quantitatively on a synthetic dataset, and qualitatively on a real dataset. The synthetic dataset consists of 100 randomly selected cases from the publicly available, multimodal IXI dataset (`brain-development.org/ixi-dataset`). After skull stripping with ROBEX [10], we generated synthetic 2D deformation fields independently for each slice, and applied them to the T2 images to simulate the geometric distortion of the histological processing. Then, we used the T1 volume as reference to recover the deformation in the T2 slices. The synthetic fields were created by composing a random similarity transformation with a nonlinear deformation; the latter was computed as the integration of a random SVF generated as smoothed Gaussian noise.

The real dataset consists of the Allen atlas [11], publicly available at http:// atlas.brain-map.org. This dataset includes a multiecho flash MRI scan acquired on a 7 T scanner at 200 µm resolution, and 679 Nissl-stained, 50 µm thick, coronal, histological sections of a left hemisphere. Manual segmentations of 862 brain structures are available for a subset of 106 sections. We downsampled the images from 1 µm to 200 µm to match the resolution of the MRI. We used the algorithm in [5] to linearly align the MRI and the stack of histological sections.

## 3.2   Experiments on Synthetic Dataset

We computed all registrations with NiftyReg, using the SVF parametrization ("vel") [9]. We affinely prealigned each distorted T2 slice to its T1 counterpart, in order to keep $\{R_k\}$ as small as possible – and hence minimize the error in the approximation in the BCH series. We computed the following registrations: 1. intermodality, between corresponding T1 and T2 slices; 2. intramodality, between each T1 slice and up to four slices right above; and 3. intramodality, between each T2 slice and up to four slices right above. The intermodality registrations used mutual information and 8 pixel control point spacing. Within modalities, we used local normalized cross correlation and 4 pixel spacing (since it is more reliable than intermodality). We then used the proposed method to recover the deformations, using between 0 and 4 intramodality neighbors – where 0 corresponds to the baseline approach, i.e., slice-wise, intermodality registration.

Figure 2a shows the root mean square (RMS) error of the registration, whereas Fig. 2b shows sample outputs of the method. The baseline approach produces very jagged contours around the cortex and ventricles. The proposed approach, on the other hand, produces smoother registrations that yield a reduction of approximately 25% in the RMS error when two neighbors are used. When a higher number of neighbors are considered, the results are still smooth, but $z$-shift starts to accumulate, leading to higher RMS error; see for instance the hypointense area on the hippocampus in the example in Fig. 2b (blue arrow).
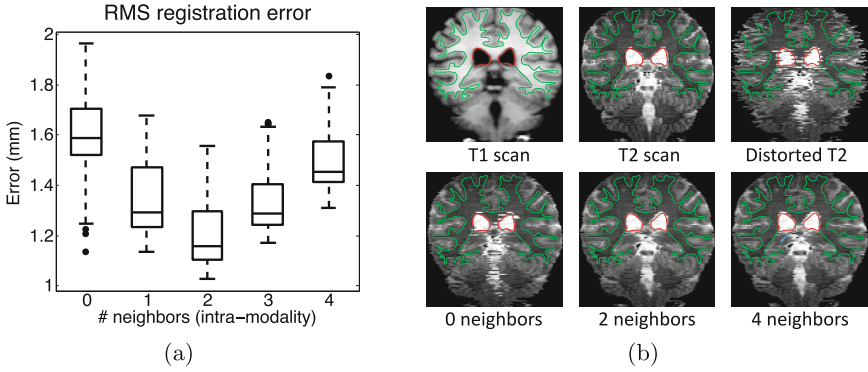
(a)                                                    (b)

**Fig. 2.** (a) Box plot for RMS registration errors in synthetic dataset. The central mark is the median; the box edges are the first and third quartile; whiskers extend to the most extreme data points not considered outliers (marked as dots). (b) A sample coronal slice from the synthetic dataset, distorted and subsequently corrected with the baseline (i.e., 0 intramodality neighbors) and proposed method (with 2 and 4 neighbors). We have superimposed the contour of the lateral ventricles (red) and white matter surface (green), manually delineated on the T1 slice. The blue arrow marks an area with $z$-shift.
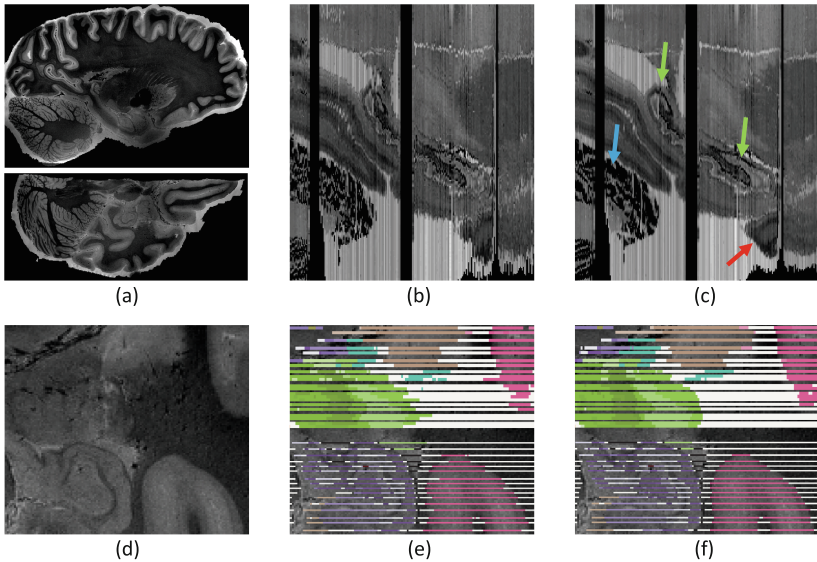


**Fig. 3.** (a) Sample slices of the *ex vivo* MRI of the Allen atlas. (b) Close-up of hippocampus, with histological reconstruction computed with the baseline approach. (c) Same region, reconstructed with the proposed method. (d) Close-up of axial MRI slice. (e) Reconstruction of manual segmentations with baseline approach. (f) Reconstruction with our method. The color map can be found on the Allen Institute website.

### 3.3   Experiments on Allen Dataset

For the Allen data, we used two neighbors within each stack, as suggested by the results on the synthetic dataset. Qualitative results are shown in Fig. 3. Much crisper reconstruction is achieved in areas such as the hippocampus (green arrows in Fig. 3c), cerebellum (blue) or cortical area 36 (red). Likewise, segmentations are more accurately propagated in areas such as the nuclei of the amygdala (different shades of green if Fig. 3e–f) and cortical regions (shades of pink).

## 4   Discussion and Conclusion

We have presented a probabilistic model for refining deformation fields in 3D histology reconstruction based on SVFs, and thus directly compatible with many widespread registration methods. Our model also serves as a starting point for future work in four main directions: 1. Inspecting better approximations to the composition of transformations; 2. Considering more realistic models for the registration errors, which account for their spatial correlation; 3. Investigating other noise models, which are more robust to outliers; and 4. Integrating an intensity homogenization model in the framework to correct for uneven section staining. The presented method is available at JEI's site: www.jeiglesias.com.

## References

1. Adler, D.H., et al.: Histology-derived volumetric annotation of the human hippocampal subfields in postmortem MRI. Neuroimage **84**, 505–523 (2014)
2. Stikov, N., et al.: In vivo histology of the myelin g-ratio with magnetic resonance imaging. Neuroimage **118**, 397–405 (2015)
3. Ourselin, S., Roche, A., Prima, S., Ayache, N.: Block matching: a general framework to improve robustness of rigid registration of medical images. In: Delp, S.L., DiGoia, A.M., Jaramaz, B. (eds.) MICCAI 2000. LNCS, vol. 1935, pp. 557–566. Springer, Heidelberg (2000). https://doi.org/10.1007/978-3-540-40899-4_57
4. Amunts, K., et al.: BigBrain: an ultrahigh-resolution 3D human brain model. Science **340**(6139), 1472–1475 (2013)
5. Malandain, G., Bardinet, E., Nelissen, K., Vanduffel, W.: Fusion of autoradiographs with an MR volume using 2-D and 3-D linear transformations. NeuroImage **23**(1), 111–127 (2004)
6. Yushkevich, P., et al.: 3D mouse brain reconstruction from histology using a coarse-to-fine approach. In: Pluim, J.P.W., Likar, B., Gerritsen, F.A. (eds.) WBIR 2006. LNCS, vol. 4057, pp. 230–237. Springer, Heidelberg (2006). https://doi.org/10.1007/11784012_28
7. Casero, R., et al.: Transformation diffusion reconstruction of 3D histology volumes from 2D image stacks. Med. Image Anal. **38**, 184–204 (2017)

8. Arsigny, V., Commowick, O., Pennec, X., Ayache, N.: A log-euclidean framework for statistics on diffeomorphisms. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 924–931. Springer, Heidelberg (2006). https://doi.org/10.1007/11866565_113

9. Modat, M., Daga, P., Cardoso, M.J., Ourselin, S., Ridgway, G.R., Ashburner, J.: Parametric non-rigid registration using a stationary velocity field. In: Mathematical Methods in Biomedical Image Analysis, pp. 145–150 (2012)

10. Iglesias, J.E., Liu, C.Y., Thompson, P.M., Tu, Z.: Robust brain extraction across datasets and comparison with publicly available methods. IEEE Trans. Med. Imaging **30**(9), 1617–1634 (2011)

11. Ding, S.L., et al.: Comprehensive cellular-resolution atlas of the adult human brain. J. Comp. Neurol. **524**(16), 3127–3481 (2016)