Synchrotron-μCT / MRI spatial registration update

Scott Trinkle

Last edited: August 24, 2018

1 Introduction

1.1 Motivation

Diffusion MRI can be used to calculate metrics that estimate the bulk orientation of sub-resolution nerve fiber microstructure, such as the tensor/ellipsoid in diffusion tensor imaging (DTI) or the fiber-orientation distribution (FOD) in high angular resolution diffusion imaging (HARDI). These orientation metrics can then be fed into tractography algorithms to estimate the connectivity of nerve bundles across the brain.

Synchrotron x-ray μCT yields images with dramatically higher spatial resolution. The high resolution allows for the detectability of individual bundles of myelinated axons. The orientation of these bundles can be estimated using structure tensor analysis, and the bulk distribution of these orientations across an arbitrary-sized ROI can be calculated, generating FODs analogous to those calculated with HARDI imaging. MRI tractography algorithms could be applied to these μCT-generated FOD algorithms and compared to the segmented nerve bundles themselves for validation. Both of these μCT based connectivity maps could also be compared to tractograms calculated from MRI data.

Nonlinear spatial registration of the μCT and MRI data is needed in order to perform quantitative, voxel-wise comparison of any of these parameter maps (FODs, tractograms, or segmented axons bundles). The lower resolution of the MRI dataset (150 μm isotropic vs. 1.2 μm isotropic for μCT) provides a challenge to this spatial registration task. The full-resolution μCT dataset is needed to accurately "ground-truth" many of these metrics, but the low resolution of the MRI dataset puts a limit on the feasible scale of quantitative validation. Additionally, the large size of the μCT dataset poses a significant computational challenge for existing registration packages / algorithms.

1.2 Diffusion MRI Validation Literature Review (focus on registration methods)

Many groups have done studies validating diffusion MRI methods with both histology [1–3] and polarized light imaging (PLI) [4].

Mollink et al. [4] compared two-dimensional dispersion and orientation metrics from diffusion MRI (400 μm isotropic resolution), histology ($0.28 \times 0.28 \times 6$ μm resolution) and PLI ($4 \times 4 \times 60$ μm resolution) in human brain samples. They calculated 2D deformable transforms using the Modality Independent Neighborhood Descriptor (MIND) algorithm [5], which emphasizes local rather than global similarity. Their input images were the mean diffusivity map from the diffusion MRI image, the I_0 map for PLI and the grayscale stained image for histology. The calculated deformation fields were applied to the 2D FOD maps, which were then reoriented using the preservation of principal directions strategy [6], in which a local affine transformation is computed for each

point in the deformation field and used to reorient the FODs. The success of the registration was assessed by measuring the distances between manually identified landmarks. In some regions, they claim accuracy to within 400 μm (i.e., one MRI voxel).

Most groups that have studied 3D [1, 2] and 2D [3] FOD validation use an MRI/histology registration methodology published by Choe et al. [7] at Vanderbilt. The data acquisition protocol is as follows:

1. Diffusion MRI is acquired at ~300-400 μm
2. The sample is sectioned at 50-80 μm. The block face is photographed every 2-3 slices. These photographs are stitched together to create a "block face volume" with through-plane resolution of 150-160 μm.
    - Choe [7] recommends downsampling the axial resolution of this block face volume, for isotropic voxels of 150-160 μm.
3. A 2D histological montage is acquired with 600-900 tiles for each slice, with nominal resolution of 0.80 μm. These tiles are stitched together with Zeiss software.
4. The Anderson group [1, 2] then identifies a number of ROI within each slice to acquire a high-resolution 3D confocal subvolume with nominal resolution of $0.18 \times 0.18 \times 0.42$ μm over a 900×900 μm field of view (equivalent to nine MRI voxels). The structure tensor FODs are calculated from these volumes.

After a number of processing/deconvolution steps to account for tissue shrinkage, nonisotropic resolution, etc., the registration pipeline proceeds as follows:

1. The relevent metrics (tensors/FODs) are calculated in their native spaces.
2. Each 2D histology montage is registered to its corresponding block face image with mutual-information based 2D linear registration and 2D nonlinear registration with the adaptive bases algorithm (ABA) [8].
3. The $b = 0$ (non diffusion-weighted) MRI image is registered to the block face volume with 3D affine registration and nonlinear transform calculated with ABA.
4. For the Anderson group, given the location of the 3D confocal ROI within the 2D montage, the combined deformation fields from MRI $\rightarrow$ block face and block face $\rightarrow$ 2D montage are used to determine the MRI signal from the same tissue volume.
    - The MRI tensors/FODs are then transformed to histological space and reoriented using the preservation of principal directions approach [6] for tensors, or the method developed by Hong et al. [9] for FODs.

Choe et al. validated this methodology by measuring the distance between 291 anatomic landmarks manually identified in both images, with mean accuracies within one MRI voxel (~300 μm).

Important to note is that with this method, the MRI is transformed to the histological space, and the structure tensor FODs from histology are calculated using MRI-voxel-sized ROI. Thus, in these studies, there is no resampling or interpolation of the final FODs from either modality. The papers do not address the multi-scale nature of their registrations from MRI $\rightarrow$ block face or block face $\rightarrow$ 2D montage.

1.3 Application to μCT validation study

This methodology can be adapted for the purposes of comparing FODs calculated from μCT and MRI data. In this case, the higher-resolution (50 μm structural) MRI data can take the place of the intermediate "block face volume," but all comparisons should be performed at the diffusion MRI scale. For FOD validation, the proposed pipeline is:

1. Calculate FODs from both datasets using their native resolutions. Structure tensor orientations will be calculated with full x-ray resolution, but the FODs will be binned over 150 μm isotropic ROI.
2. Downsample the μCT data and structural MRI data to the diffusion MRI resolution.
3. Calculate the linear and nonlinear μCT → structural-MRI and linear structural-MRI → diffusion-MRI spatial transformations.[1]
4. Apply these transformations to the μCT FODs (already binned at diffusion MRI resolution) to get them into the diffusion MRI space
5. Reorient the FODs appropriately using the methods described in [6, 9].
6. Perform voxel-wise quantitative comparisons (peak angular error, angular correlation coefficient, Jensen-Shannon divergence, etc.)

For tractography validation, the same pipeline will apply. If μCT tractography is performed on registered MRI-scale FODs, then no further work is needed; the same algorithm can be applied to both datasets and any number of common metrics can be used to evaluate the results [10, 11]. Tractograms can be calculated with subvoxel resolution, so to compare segmented axon bundles from the μCT to tractography results from either dataset, the difference would appear in Step 2: the μCT and structural MRI data will be downsampled to the tractogram resolution instead of the diffusion MRI data resolution. New transforms will then be calculated at that resolution and applied to the segmentation results.

1.4 Aim

The original aim of this report was to detail the methods involved in a different pipeline and to characterize its performance as a function of scale and interpolation method. This original pipeline involved calculating the registration at an intermediate resolution achieved by upsampling the MRI data. I no longer think this is a fair or useful approach. The goal of this study is to validate commonly used methods of estimating orientation and connectivity information from diffusion MRI data. The idea of upsampling the MRI data to an intermediate resolution was to avoid "wasting" detail present in the μCT data — but we never expect diffusion MRI to be able to report that detail in the first place. Assuming these MRI methods are accurate, we expect the results to be a downsampled, bulk representation of the "true" results calculated from the μCT data. Performing the validation at the native MRI scale will allow us to fairly evaluate that assumption.

The bulk of an earlier iteration of this report involved a comparison between two strategies for calculating the transform at an intermediate scale. With the new proposed pipeline, this question is less relevant — but for clarity, the strategies are:

[1] Like Choe et al. and the Anderson group, we could assess registration accuracy using the distance between manually identified landmarks. We could either do this ourselves or find a mouse brain anatomist...

1. "Resampled Data" — both μCT and MRI data are scaled to some common resolution x, and the transformation is calculated and applied.

2. "Resampled Transform" — the μCT data is scaled to the native MRI resolution, the transform is calculated, then the transform itself is scaled to some higher resolution x and applied.

We expected the "Resampled Data" approach to achieve more quantitatively accurate registrations, but thought the "Resampled Transform" approach might be preferable in situations where it becomes too computationally expensive to calculate the registration at the desired resolution.

The remainder of this report is unchanged.

2 Methods

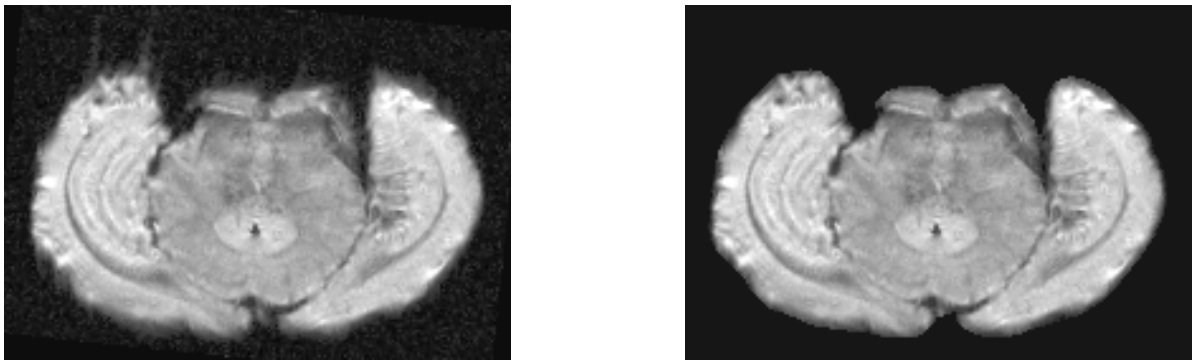
2.1 μCT downsampling

The full 32-bit μCT data has a resolution of 1.2 μm isotropic. For this study, an 8-bit, 4x downsampled (4.8 μm isotropic) single-file volume image was taken as the "initial" μCT data. This was the highest resolution version of the μCT data that existed as a single file of manageable size. This data was then downsampled to voxel sizes of 15, 20, 25, 40, 50, 60, 75, 100, and 150 μm isotropic using bicubic interpolation in ImageJ.

2.2 MRI Resampling

There are two separate MRI images: a 150 μm isotropic image with 30 diffusion-weighting directions, and a 50 μm isotropic T_2-weighted (?) structural image. All registrations were calculated using the structural image, since it was assumed that its higher resolution would lead to better performance. There is a known affine transformation between these two images that can be applied to metrics calculated from the diffusion image in the future.

A mask of the structural image was calculated using a median-otsu thresholding method in the Dipy python package [12] and applied to eliminate edge distortions and background noise.[2]



(a) Raw data (b) Masked data

Figure 1: Comparison of raw and masked structural MRI data. Note the undefined boundary in the raw image.

[2]This masking was performed for a separate/previous study, but I thought it would improve registration performance to have a harder boundary to the brain volume compared to the "feathery" boundary of the raw data. See Figure 1.

The masked structural image was then resampled to the same resolutions as the μCT data: 15, 20, 25, 40, 50, 60, 75, 100, and 150 μm isotropic, using bicubic, bilinear and "no" interpolation in ImageJ.[3]

2.3 Diffeomorphic registration

Diffeomorphic registration was performed for every resolution and interpolation method using the Advanced Normalization Tools (ANTs) software package, a interface for the Insight ToolKit image registration framework [13]. ANTs primarily uses the NIfTI-1 (Neuroimaging Informatics Technology Initiative) data format for image input/output. All images were converted to this format and the physical voxel dimensions were manually added to the metadata.

Both affine and diffeomorphic transformations were calculated using the default parameters of the `antsRegistrationQuickSyN.sh` script. Details are available here, but a quick summary is provided below:

1. A rigid transformation is calculated to align the centers of mass of the two images. This is performed in four stages, with downsampling factors of 12, 8, 4 and 2, smoothing sigmas of 4, 3, 2 and 1 and maximum iteration numbers of 1000, 500, 250 and 100, respectively. The mutual information (with 32 bins) is optimized in each stage until successive iterations have a difference less than 10^{-6} or until the maximum number of iterations is reached.

2. An affine transformation is calculated using the same parameters as above.

3. A nonlinear, diffeomorphic transformation field is calculated with the competition-winning [14] ANTs Symmetric Normalization (SyN) algorithm [15]. This is performed in five stages, with downsampling factors of 10, 6, 4, 2, and 1, smoothing sigmas of 5, 3, 2, 1 and 0 and maximum iteration numbers of 100, 100, 70, 50 and 20, respectively. The cross correlation metric is optimized with a radius of 4, to the same convergence criteria as above.

The output is the composite linear transform (4x4 matrix), transform field (dimension: $X \times Y \times Z \times 3$ for image dimension $X \times Y \times Z$), and the final transformed image. For these registrations, the μCT image was taken as the "moving" image and transformed to the "fixed" structural MRI space. This is to avoid the eventual complication of performing nonlinear transformations on orientation-encoded diffusion MRI datasets; however, the diffeomorphic SyN transformation calculated by ANTs is invertible.

2.4 Transform resampling

The transform field calculated at 50 μm (native structural MRI resolution) was resampled to each of the other resolutions using each of the three interpolation methods. Both the composite linear transformation and the resampled transform field from the 50 μm registration were then applied to each of the μCT datasets.

Accordingly, there were a total of six transform fields calculated at each resolution: three calculated by the "Resampled Data" method and three by the "Resampled Transform" method.

2.5 Analysis

The mean squared error (MSE), Pearson product-moment correlation coefficient, and mutual information (MI) were used to quantify the similarity between various transformations as well as final transformed images.

[3] I assume that the "Interpolation: None" option in ImageJ indicates nearest neighbor interpolation

For N elements, the MSE was calculated as

$$\mathrm{MSE}(X;Y) = \frac{1}{N}\sum_{n=1}^{N}(X_n - Y_n)^2\,, \tag{1}$$

for objects (images or transforms) X and Y. A low MSE indicates high similarity between objects.

The correlation coefficient was determined by calculating a linear fit between corresponding values in two objects. In Python, this was implemented with the `corrcoef` function in the NumPy library, with flattened arrays of the two objects as inputs. A correlation coefficient close to 1 indicates high similarity between objects.

Mutual information is defined as

$$\mathrm{MI}(X;Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\ln\left(\frac{p(x,y)}{p(x)p(y)}\right), \tag{2}$$

where $p(x,y)$ is the joint probability function of X and Y, and $p(x)$ and $p(y)$ are the marginal probabilties of X and Y, respectively. A high mutual information indicates high similarity between objects. The $p(x,y)$ was implemented using the normalized NumPy `histogram2d` function with 1024 bins. The marginal probabilities were calculated by summing $p(x,y)$ over each axis (equivalent to calculating two 1024-bin one-dimensional histograms).

3 Results

3.1 Affine Transform

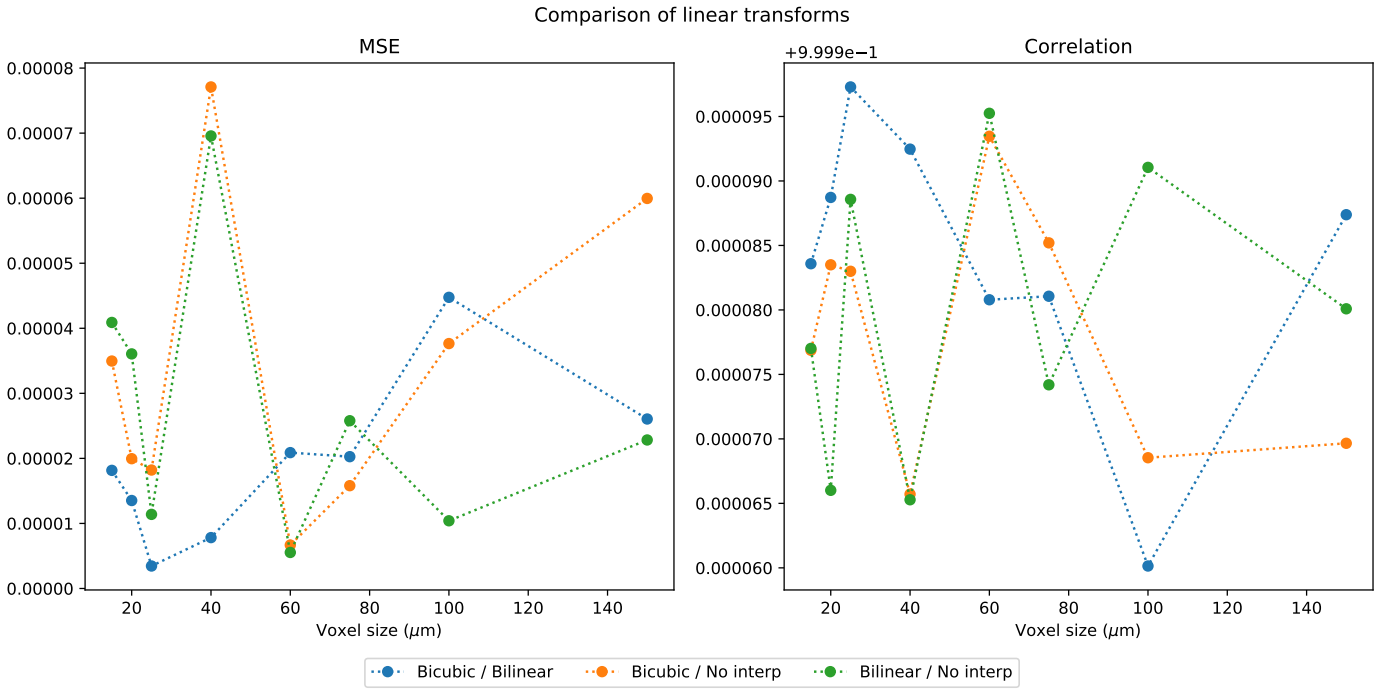


Figure 2: Comparison of linear transform elements. There are no clear patterns or trends with respect to resolution or interpolation method.

Figure 2 shows the MSE and correlation of the 12 linear transform elements as a function of interpolation

and resolution. Accordingly, this figure does not assess the accuracy of the transforms, but their stability across different interpolation scales and methods. Overall, the transforms show extremely good agreement with each other, with very low MSE and high correlation across all resolutions and interpolation methods.

3.2 Transform Field

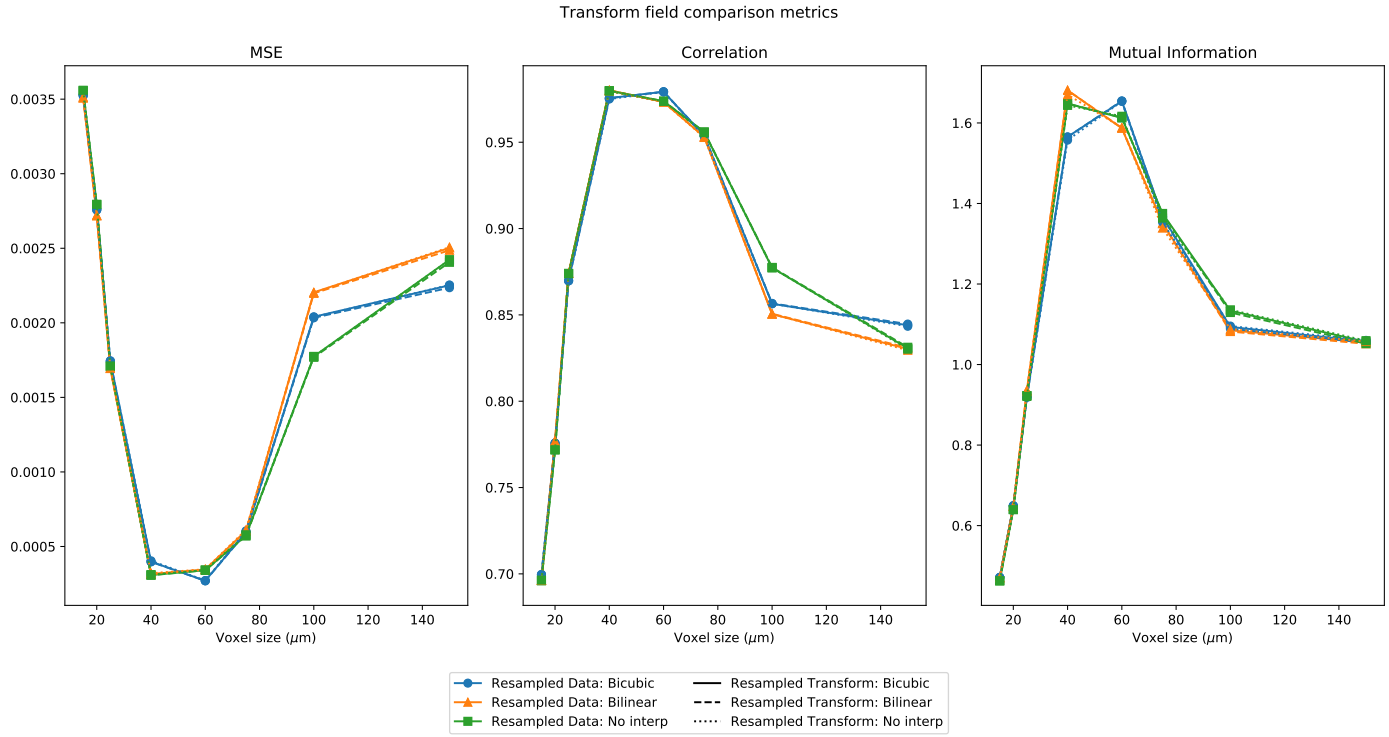


Figure 3: Comparison of transform fields calculated from resampled data to those calculated at the native resolution and then resampled.

Figure 3 compares the transform fields calculated from resampled MRI data to those calculated at 50 μm and resampled to the various resolutions. There are nine curves plotted for each metric: each of the three "Resampled Transform" interpolation methods is compared to each of the three "Resampled Data" methods for each resolution. Line color and marker style indicate the "Resampled Data" interpolation method in the comparison, and the line style indicates the "Resampled Transform" interpolation method (for example, the solid, orange line with triangle markers is plotting the comparison between the transform calculated with bilinear-interpolated MRI data and the bicubic-interpolated transform originally calculated with native-resolution MRI data).

There are a few notable trends in these plots. First, curves with different line styles are virtually identical, indicating that the similarity of the transforms depends much more on the data interpolation method than the transform interpolation method. In other words: for a given resolution, there is more variability between transforms calculated *at* that resolution using differently interpolated data than there is between a single calculated transform undergoing different interpolations *to* that resolution.

There is also a peak similarity associated with each metric around the native resolution of 50 μm. This leads to the unsurprising result that the "resampled data" and "resampled transform" approaches are the most similar near the native resolution at which the resampled transform is calculated. In other words, as you get further from 50 μm resolution, resampling the 50 μm transform will yield an increasingly worse estimate of the

transform calculated from resampled data, with very little dependence on interpolation method. Notably, this effect is more severe at smaller voxel sizes, where there is much more μCT data to utilize when calculating the registration.

3.3 Registration performance

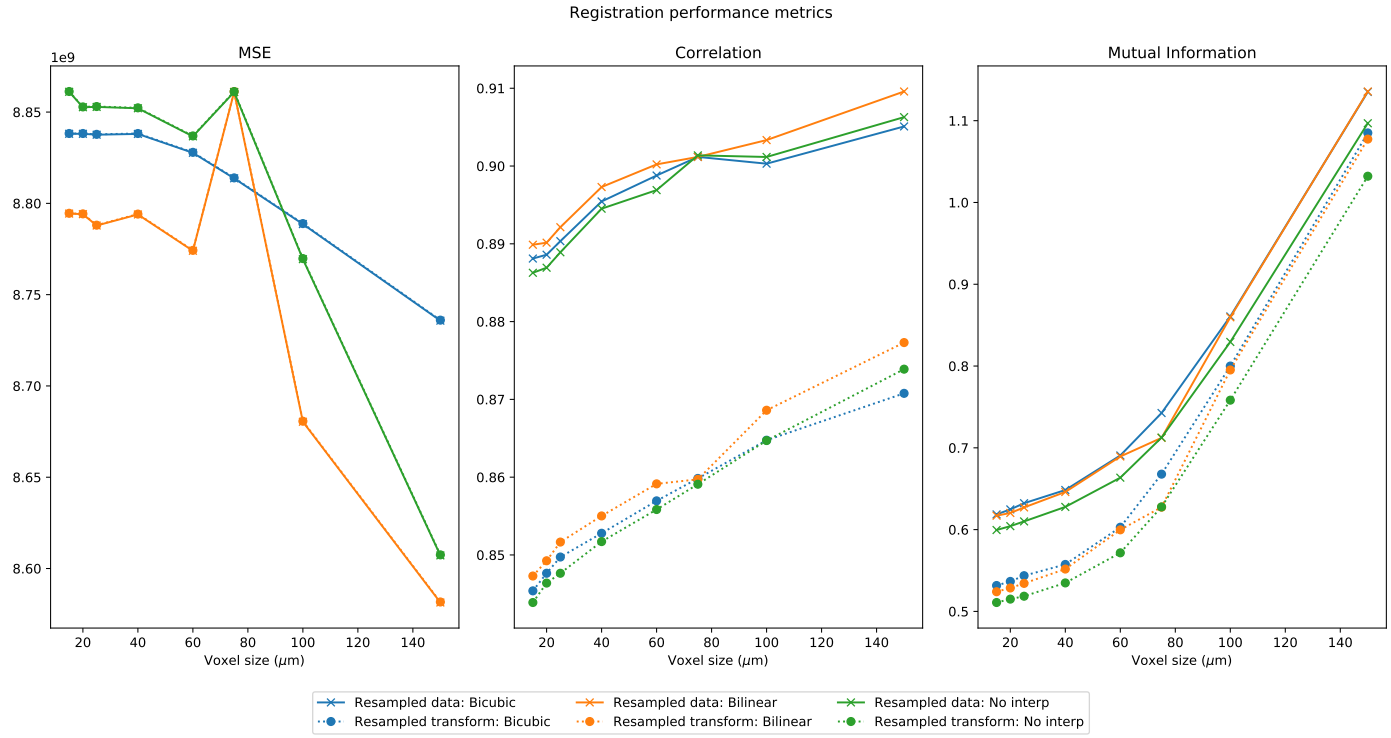


Figure 4: Comparison of transformed μCT data to MRI from all methods of calculating the transform.

The results of Figure 3 do not necessarily have implications on which of the two methods is preferable, just that they diverge as you move further from the native resolution. Figure 4 plots the similarity between the interpolated MRI data and the transformed μCT data. The line and marker styles indicate either "resampled data" or "resampled transform" approaches. The color indicates the interpolation method used in both cases. The interpolation method used to either calculate or resample the transform that was applied to the μCT data always matched the interpolation method used to create the MRI data in the comparison.

One immediate concern is the dramatically high MSE value at all resolutions (note the 1e9 indicator at the top of the axis). This results from the fact that the input μCT data are 8-bit [0,255] integer images, while the input MRI data are 32-bit images with means of ~50,000. Since this effect was common to all registrations/transforms, the relative trend is still noteworthy; however, notice that it is difficult to visualize the difference between the "resampled data" and "resampled transform" curves using the MSE metric.

In general, the main trends are as follows:

- Similarity metrics are "worse" at higher resolutions. This is expected: at higher resolutions, there is an increase in μCT information, but no corresponding increase in MRI information. It becomes more and more challenging to align small, detailed structures in the μCT data when those structures are missing entirely in the MRI data.

- The "resampled data" approach (solid lines) yields "better" similarity metrics for every resolution and interpolation metric. This is expected as well: transformations calculated from interpolated data are able to be fine-tuned to match the deviations / blurring introduced in interpolation. From the mutual information metric, we see that this deviation between the two methods grows larger at higher resolutions.
- The "optimal" interpolation metric depends on resolution and changes with different similarity metrics. Generally, bilinear or bicubic interpolation perform the best, and no interpolation performs the worst.

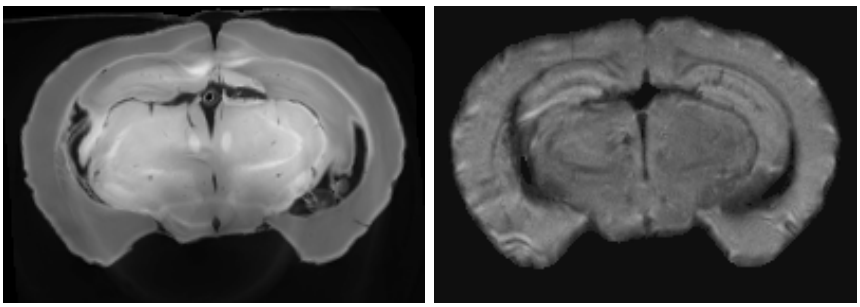
4 Conclusions

As expected, it is clear that there is a difference between the "resampled data" and "resampled transform" approaches. It is not immediately clear whether or not the results in Figure 4 indicate that one method is preferable to another. Resampling the data before calculating the transform yields a more quantitatively accurate registration, but the useful information at any resolution is limited to that of the original image — interpolating to a higher resolution cannot generate new information.

This point is illustrated in Figures 5 and 6. Figure 5 shows the results of a transformation calculated at the native MRI resolution of 50 µm. In Figure 6a, this "native" transform is upsampled to 15 µm and applied to the µCT data. In contrast, Figure 6b shows the transformation calculated using MRI data upsampled to 15 µm (Figure 6c). The "resampled data" result is quantiatively more similar to Figure 6c than the "resampled transform" result is, but Figure 6a appears more realistic and less warped; note the wavy boundary in Figure 6b.

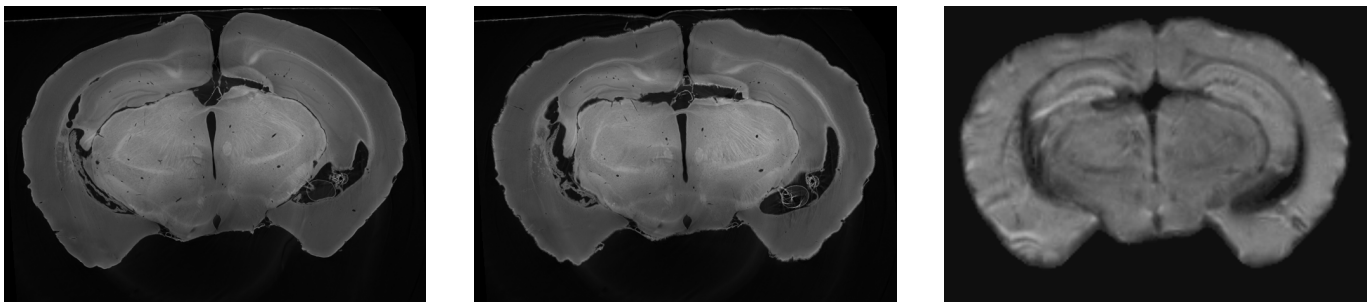
In terms of applying these transformations to calculated metrics (FODs, etc.), I believe that quantitative registration accuracy will be preferred over "realistic"-looking images when it comes to voxel-wise comparisons. For this reason, the "resampled data" approach is most likely preferred, provided it is computationally feasible at the desired scale. We want the registration to account for the changes introduced by interpolation.

However, we need to think about whether or not we actually have anything to gain by upsampling the MRI metrics at all, particularly if the final goal is to validate the MRI techniques used to calculate them. As Figure 4 shows, we only expect the registration performance (and thus, the accuracy of future metric comparisons) to decrease at higher resolutions: we are not interested in the ability of *upsampled* diffusion MR to report connectivity information. It might be best to just perform all comparisons at the native MRI resolution.



(a) µCT (b) MRI

Figure 5: Transformation calculated at native MRI resolution (50 µm).



(a) μCT: "resampled transform" (b) μCT: "resampled data" (c) MRI

Figure 6: Transformations calculated with either "resampled data" or "resampled transform" approaches (15 μm).

5 Future work

In preparing this report, I came across a few changes I would make to future registrations. I do not suspect that all of these would have a dramatic effect on the comparative results presented in Figures 2-4, but they might improve registration accuracy:

- Mask μCT data as well as MRI data
- Normalize the means of two images before registration
- Use MI instead of CC as optimization metric in ANTs SyN algorithm
  - This metric is commonly preferred for images with different contrast
- Let ANTs run to convergence for all registrations
  - It reached the max iteration limit for large images, which might play a role in the lower reported performance metrics at these scales.

Additionally, I will look closer at the multi-modality, multi-scale registration literature to see if there are any other insights I have looked over.

References

[1] K. Schilling, V. Janve, Y. Gao, I. Stepniewska, B. A. Landman, and A. W. Anderson, "Comparison of 3D orientation distribution functions measured with confocal microscopy and diffusion MRI," *NeuroImage*, vol. 129, pp. 185–197, 2016.

[2] K. G. Schilling, V. Janve, Y. Gao, I. Stepniewska, B. A. Landman, and A. W. Anderson, "Histological validation of diffusion MRI fiber orientation distributions and dispersion," *NeuroImage*, vol. 165, pp. 200–221, jan 2018.

[3] A. Seehaus, A. Roebroeck, M. Bastiani, L. Fonseca, H. Bratzke, N. Lori, A. Vilanova, R. Goebel, and R. Galuske, "Histological validation of high-resolution DTI in human post mortem tissue," *Frontiers in Neuroanatomy*, vol. 9, no. July, pp. 1–12, 2015.

[4] J. Mollink, M. Kleinnijenhuis, A. M. van Cappellen van Walsum, S. N. Sotiropoulos, M. Cottaar, C. Mirfin, M. P. Heinrich, M. Jenkinson, M. Pallebage-Gamarallage, O. Ansorge, S. Jbabdi, and K. L. Miller, "Evaluating fibre orientation dispersion in white matter: Comparison of diffusion MRI, histology and polarized light imaging," *NeuroImage*, vol. 157, no. December 2016, pp. 561–574, 2017.

[5] M. P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. V. Gleeson, S. M. Brady, and J. A. Schnabel, "MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration," *Medical Image Analysis*, vol. 16, pp. 1423–1435, oct 2012.

[6] D. C. Alexander, C. Pierpaoli, P. J. Basser, and J. C. Gee, "Spatial transformations of diffusion tensor magnetic resonance images," *IEEE Trans Med Imaging*, vol. 20, no. 11, pp. 1131–1139, 2001.

[7] A. S. Choe, Y. Gao, X. Li, K. B. Compton, I. Stepniewska, and A. W. Anderson, "Accuracy of image registration between MRI and light microscopy in the ex vivo brain," *Magnetic Resonance Imaging*, vol. 29, no. 5, pp. 683–692, 2011.

[8] G. Rohde, A. Aldroubi, and B. Dawant, "The adaptive bases algorithm for intensity-based nonrigid image registration," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 1470–1479, nov 2003.

[9] X. Hong, L. R. Arlinghaus, and A. W. Anderson, "Spatial normalization of the fiber orientation distribution based on high angular resolution diffusion imaging data," *Magnetic Resonance in Medicine*, vol. 61, pp. 1520–1527, jun 2009.

[10] L. Ning, F. Laun, Y. Gur, E. V. DiBella, S. Deslauriers-Gauthier, T. Megherbi, A. Ghosh, M. Zucchelli, G. Menegaz, R. Fick, S. St-Jean, M. Paquette, R. Aranda, M. Descoteaux, R. Deriche, L. O'Donnell, and Y. Rathi, "Sparse Reconstruction Challenge for diffusion MRI: Validation on a physical phantom to determine which acquisition scheme and analysis method to use?," *Medical Image Analysis*, vol. 26, pp. 316–331, dec 2015.

[11] C. Thomas, F. Q. Ye, M. O. Irfanoglu, P. Modi, K. S. Saleem, D. A. Leopold, and C. Pierpaoli, "Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited," *Proceedings of the National Academy of Sciences*, vol. 111, no. 46, pp. 16574–16579, 2014.

[12] E. Garyfallidis, M. Brett, B. Amirbekian, A. Rokem, S. Van Der Walt, M. Descoteaux, and I. Nimmo-Smith, "Dipy, a library for the analysis of diffusion mri data," *Frontiers in Neuroinformatics*, vol. 8, p. 8, 2014.

[13] B. B. Avants, N. J. Tustison, M. Stauffer, G. Song, B. Wu, and J. C. Gee, "The insight toolkit image registration framework," *Frontiers in Neuroinformatics*, vol. 8, p. 44, 2014.

[14] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration," *NeuroImage*, vol. 46, no. 3, pp. 786 – 802, 2009.

[15] B. Avants, C. Epstein, M. Grossman, and J. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Medical Image*

Analysis, vol. 12, no. 1, pp. 26 – 41, 2008. Special Issue on The Third International Workshop on Biomedical Image Registration – WBIR 2006.