Combining Digital Histology with Magnetic Resonance Imaging in Amyotrophic Lateral Sclerosis

Istvan N Huszar

Kellogg College
University of Oxford

Submitted in partial fulfilment of the degree
Doctor of Philosophy

Michaelmas 2020

Abstract

There is an increasing demand for non-invasive biomarkers that can reliably indicate the presence and extent of neurodegeneration in its early stages. Various magnetic resonance imaging (MRI) methods have shown great potential to indicate the event of upper motor neuron involvement in the progression of amyotrophic lateral sclerosis (ALS) at the population level, but almost none of these have been systematically validated against traditional histological markers at the individual level. Validation relies on a precise and preferably automatic method to align a large number of MRI and histology images of the same tissue, which poses unique challenges compared to more conventional MRI-to-MRI registration.

The thesis summarises the development and evaluation of a novel image registration framework (Tensor Image Registration Library) that is designed to achieve automated registration between sparsely sampled histology sections and whole-brain post-mortem MR volumes derived from an ALS study cohort. A multi-stage automated pipeline is developed, in which two types of histology sections are registered to an MRI volume via one or two photographic intermediates. The research chapters explain the rationale behind each stage, describe the algorithm, and report the experiments that were carried out to verify the accuracy and robustness of the implementation.

The proposed method uniquely integrates part-to-whole matching and 2D-to-3D registration, allowing conventional small histology samples to be registered without serial sectioning. Furthermore, the orthogonal deformations of a slice are estimated to account for the imperfections of a free-hand cut, which are shown to be as large as several millimetres. The pipeline achieves sub-millimetre accuracy in each stage and equally accurate alignment between histology and MRI when the stages are combined. This registration approach alleviates the need for specialist cutting or stain automation hardware, and is therefore suitable for integration into a wide range of studies that utilise routine neuropathology practices.

Combining Digital Histology with Magnetic Resonance Imaging in Amyotrophic Lateral Sclerosis



Istvan N Huszar

Kellogg College

University of Oxford

Submitted in partial fulfilment of the degree

Doctor of Philosophy

Michaelmas 2020

Acknowledgements

First of all, I would like to express my sincere gratitude to the *Clarendon Fund* and the *Chadwyck-Healey Charitable Trust* at *Kellogg College* for giving me the opportunity to spend four wonderful years at Oxford that were altogether a truly transformative experience for me.

Further, I would like to thank my supervisors *Karla Miller*, *Mark Jenkinson* and *Peter Jezzard* for taking me on this exciting journey. Your enthusiasm and immense knowledge have always been a source of my inspiration for more. I was a medic who came unsuspecting to the field of engineering and your empathy and professionalism paved my way to become an all-round open-minded academic. I cannot thank you enough for the countless times when you gave me thoughtful comments on this work; your guidance has helped me greatly to finish my dissertation.

I am also thankful to the members and staff of the *MPLS Doctoral Training Centre* and the *Nuffield Department of Clinical Neurosciences* for hosting me through these years, as I thoroughly enjoyed working with you all. I am especially thankful to *Olaf Ansorge*, *Martin Turner*, *Menuka Pallebage-Gamarallage*, *Evan Edmond*, *Jeroen Mollink*, *Daniel Papp*, *Mark Chiew*, *Benjamin Tendler*, and *Wenchuan Wu* for their help with my research projects. I will not forget the kindness of *Paul McCarthy* and *Matthew Webster* to help me release my work to the public. Special thanks to my ONBI-16 friends: *Amy*, *Rick* and *Sophie*, and fellow DPhil students *Chaoyue*, *Daniel*, *Sam* and *Feri* for the inspiring discussions from time to time.

I am very grateful to the help and continuous encouragement of my professional friends: *Miklós Kellermayer*, *Péter Sótonyi*, and *Miklós Garai*, and to the people who were always close to me even when we were physically far: *Laci*, *Dara*, *Soma*, *Eszter*, *Kinga*, *Dorka*, *Rozina* and *Siyang*.

I cannot thank enough my family for their continued support and faith in me. *Csilla*, you know that I would have never done this all by myself. We did this together and I will always be grateful that you have helped me pursue my dreams. *Mama*, I know that you have always wanted me to become a doctor... I finally have good news for you: I'll hopefully do it now for the second time, thanks to your love and faith in my decisions. Of course, I cannot forget to thank the love and support of the people who now live in my remembrance.

Finally, I am humbled by the generous donation of the ALS patients who have contributed to this work, and I express my gratitude to their families for keeping their faith in us. This may be just a small step, but we will one day find a cure for ALS.

Abstract

There is an increasing demand for non-invasive biomarkers that can reliably indicate the presence and extent of neurodegeneration in its early stages. Various magnetic resonance imaging (MRI) methods have shown great potential to indicate the event of upper motor neuron involvement in the progression of amyotrophic lateral sclerosis (ALS) at the population level, but almost none of these have been systematically validated against traditional histological markers at the individual level. Validation relies on a precise and preferably automatic method to align a large number of MRI and histology images of the same tissue, which poses unique challenges compared to more conventional MRI-to-MRI registration.

The thesis summarises the development and evaluation of a novel image registration framework (Tensor Image Registration Library) that is designed to achieve automated registration between sparsely sampled histology sections and whole-brain post-mortem MR volumes derived from an ALS study cohort. A multi-stage automated pipeline is developed, in which two types of histology sections are registered to an MRI volume via one or two photographic intermediates. The research chapters explain the rationale behind each stage, describe the algorithm, and report the experiments that were carried out to verify the accuracy and robustness of the implementation.

The proposed method uniquely integrates part-to-whole matching and 2D-to-3D registration, allowing conventional small histology samples to be registered without serial sectioning. Furthermore, the orthogonal deformations of a slice are estimated to account for the imperfections of a free-hand cut, which are shown to be as large as several millimetres. The pipeline achieves sub-millimetre accuracy in each stage and equally accurate alignment between histology and MRI when the stages are combined. This registration approach alleviates the need for specialist cutting or stain automation hardware, and is therefore suitable for integration into a wide range of studies that utilise routine neuropathology practices.

Contents

List of Figures viii

List of Tables xviii

List of Abbreviations xx

1 Introduction 1
- 1.1 Amyotrophic Lateral Sclerosis . . . 1
- 1.2 Clinical Imaging in ALS . . . 2
- 1.3 Post-mortem Imaging in ALS . . . 2
- 1.4 Combining Histology and MRI Data in ALS . . . 3
- 1.5 Organisation of the thesis . . . 4

2 Medical Image Registration 6
- 2.1 Definition . . . 6
  - 2.1.1 Scope . . . 6
  - 2.1.2 Formulation as a mathematical optimisation problem . . . 7
- 2.2 Types of input images . . . 8
- 2.3 Transformation models . . . 8
  - 2.3.1 Linear . . . 8
  - 2.3.2 Non-linear . . . 10
- 2.4 Similarity metrics . . . 11
  - 2.4.1 Landmark-based metrics . . . 11
  - 2.4.2 Intensity-based metrics . . . 12
- 2.5 Modality-independent neighbourhood descriptor . . . 13
- 2.6 Optimisation . . . 16
  - 2.6.1 Gradient-based methods . . . 17
  - 2.6.2 Gradient-free methods . . . 18
  - 2.6.3 Discrete methods . . . 19
  - 2.6.4 Heuristic methods . . . 19
- 2.7 Previous work in Histology-to-MRI registration . . . 20
  - 2.7.1 Volume-based methods . . . 20
  - 2.7.2 Single-slice methods . . . 21
  - 2.7.3 Considerations for a new image registration framework . . . 23

3 Stage 1: Histology-to-block registration **24**
3.1 Dataset . . . 25
3.1.1 Post-Mortem Magnetic Resonance Imaging . . . 25
3.1.2 Dissection photographs . . . 26
3.1.3 Histology . . . 27
3.2 Histology-to-MRI Registration in three stages . . . 27
3.3 Rationale . . . 28
3.4 Distortions and artefacts of the input images . . . 29
3.4.1 Histology . . . 29
3.4.2 Tissue block photograph . . . 30
3.5 Image pre-processing . . . 31
3.5.1 Tissue-block photograph . . . 31
3.5.2 Histology . . . 32
3.6 Registration approach . . . 33
3.6.1 Transforming the images to physical space . . . 34
3.6.2 Deconstructing the transformation chain of the tissue block . 36
3.6.3 Deconstructing the transformation chain of the histology image 36
3.6.4 Pixel-to-pixel mapping by the unified stage-1 transformation chain . . . 40
3.6.5 Optimisation substages . . . 40
3.7 Evaluating the accuracy of the registration . . . 47
3.8 Exploring the effects of regularisation . . . 49
3.9 Registering sections with different stains . . . 54
3.10 A comparison with ANTs SyN . . . 58

4 Stage 2: Block-to-slice registration **61**
4.1 Rationale . . . 62
4.2 Image pre-processing . . . 63
4.2.1 Tissue block photographs . . . 63
4.2.2 Brain slice photographs . . . 63
4.3 Automating block initialisation . . . 65
4.4 Registration approach . . . 67
4.5 Robustness of finding insertion sites . . . 70
4.6 Initialisation of competing blocks . . . 73
4.7 Results of block-to-slice registration . . . 77
4.8 Effect of block initialisation on the registration accuracy . . . 79

5 Stage 3: Slice-to-volume registration 83
5.1 Rationale . . . 84
5.2 Image pre-processing . . . 86
5.2.1 Brain slice photograph . . . 86
5.2.2 MRI volume . . . 86
5.3 Transformation chain and brain slice initialisation . . . 88
5.4 Registration approach . . . 92
5.4.1 Substage 1: rigid search . . . 92
5.4.2 Substage 2: affine alignment . . . 93
5.4.3 Substage 3: in-plane deformation . . . 94
5.4.4 Substage 4: free-form deformation . . . 95
5.4.5 Substage 5: affine re-estimation . . . 96
5.5 Measuring the registration accuracy on simulated slices . . . 97
5.6 Evaluating the MIND metric versus mutual information . . . 101
5.7 Control point optimisation . . . 104
5.7.1 Control point placement schemes . . . 104
5.7.2 Number of control points . . . 110
5.8 Registration of a real brain slice photograph . . . 114
5.9 Comparison with manual landmarks . . . 116

6 Histology-to-MRI Registration Experiments with TIRL 119
6.1 Registration by two photo intermediates . . . 120
6.1.1 Dataset . . . 120
6.1.2 Methods . . . 122
6.1.3 Results . . . 129
6.1.4 Discussion . . . 149
6.2 Registration by one photo intermediate . . . 150
6.2.1 Dataset . . . 150
6.2.2 Methods . . . 151
6.2.3 Result . . . 155
6.2.4 Discussion . . . 156

7 Concluding remarks 159
7.1 Thesis contributions . . . 159
7.2 Future work . . . 161
7.2.1 Scientific and technical developments of TIRL . . . 161
7.2.2 Registration without photographic intermediates . . . 162
7.2.3 Histology-informed biomarkers . . . 164
7.3 Conclusion . . . 167

Appendices

A Tensor Image Registration Library **169**
A.1 What is TIRL? 169
A.2 What are the components in TIRL? 170
A.3 What makes TIRL a unique framework? 172

References **174**

List of Figures

2.1 **(A) Calculation of the MIND vector** at $\mathbf{x_0}$ is based on the intensity values of the pixel at $\mathbf{x_0}$ and in its N-element neighbourhood (*red tiles*). In the simplest and most common case, a pixel's neighbourhood is defined by the immediate 8 neighbours in 2-D (or 26 neighbours in 3-D). Alternative definitions, excluding the diagonal elements or using different parametrisations are also possible. The schematic shows a hollow-plate neighbourhood parametrised by its inner (R_i) and outer (R_o) half-side lengths to avoid the clutter and provide greater clarity over the calculation steps. Pixel intensity relationships are calculated between image patches as opposed to individual pixels for robustness. A central patch (*blue*) of 9 pixels (or 27 voxels) is defined around $\mathbf{x_0}$, and N neighbourhood patches (*green*) of the same size are defined around the N pixels of the neighbourhood. The i-th component of the MIND vector (m_i) represents standardised *patch-wise distances* (i.e. variance-normalised sum of squared intensity differences between corresponding pixel intensities) of the central patch and the i-th neighbourhood patch. (The patch-wise distances are calculated via Gaussian convolution, hence the additional σ parameter, which effectively represents the patch size.) **(B) MIND representation of a histology image** (the image was converted to grayscale before the operation). Black-and-white images on the periphery depict the individual components of the MIND vectors (calculated for all pixels in the sample). The position of the tiles is directly related to the position of the neighbourhood patch that was used in the calculation of the respective component. Note the direction-sensitive enhancement of image edges in each tile. 14

3.1 Overview of the post-mortem biomarker study data collection pipeline. 25

3.2 Overview of the proposed three-stage histology-MRI registration pipeline. The two photographic intermediates help with the automation of the pipeline. 28

3.3 Segmenting tissue blocks from the original autopsy photographs. . . 31

3.4 Diagram of the two TIRL transformation chains that are used in stage 1 to align the histology image (H) with the tissue block photograph (B). *Black hexagons* D_0 and D_5 represent the pixel domains of the images that the transformations are applied to. (These correspond to actual `Domain` objects in the implementation of stage 1.) D_3 is the domain of the non-linear transformation in physical space, which is instantiated from D_0 with the internal, init and linear subchains. *White hexagons* S_1 and S_2 are virtual entities, representing the intermediate, partially transformed histology domains that stand as the source and target domains of the init subchain. S_4 (*blue hexagon*) denotes the (virtual) common target of the H and B chains. 34

3.5 Large orientation differences cannot be reliably compensated by local optimisation, as the state associated with the lowest cost is only accessible through states that are associated with higher-than-initial cost. 41

3.6 Stage-1 accuracy is evaluated by mean contour distance. Contours (*red*, *blue*) are hand-drawn at the gray-white matter interface on the tissue block photograph **(A)** and the histology image **(B)**. **(C–F)** show a gradually improving overlap between the contours after each substage of the registration (see main text). The *yellow arrowheads* indicate the largest changes relative to the previous substage. As indicated by the mean contour distances, the non-linear stage is the largest contributor to registration accuracy. Note: only the contours are mapped onto the histology image – the image itself is identical in all panels of the figure. 49

3.7 Non-linear registration of a hippocampal section after affine initialisation using varying degrees of regularisation ($0.2 \leq \alpha \leq 1.0$). All images are show on the domain of the original tissue block photograph (D_5). The histology image in the top row is shown after the affine substage was completed. *Red grid lines* are provided for comparison, *gray grid lines* show the estimated deformations of the histology image. Jacobian ranges (*jac*) and mean contour distances (*acc*) are reported for each case. The *blue contour* was obtained by transforming the contour from the original histology image, the *red contour* was drawn on the original tissue block photograph to highlight the accuracy of the alignment. 50

3.8 Accuracy of stage-1 registration measured by mean contour distance on 14 callosal sections with manually-defined ground truth. Accuracy is reported separately for each linear substage (*white background*), as well as the non-linear substage, using different values of the regularisation parameter α (*green background*). Horizontal lines represent the median, boxes show the interquartile range, and the bars span the whole range of the data. 52
3.9 Accuracy of stage-1 registration measured by mean contour distance on 14 hippocampal sections with manually-defined ground truth. Accuracy is reported separately for each linear substage (*white background*), as well as the non-linear substage, using different values of the regularisation parameter α (*green backgorund*). Horizontal lines represent the median, boxes show the interquartile range, and the bars span the whole range of the data. 53
3.10 Stage-1 results on a section of the orbitofrontal cortex immuno-stained for PLP, Ferritin and tincture-stained with LFB+PAS. All these stains exhibit good gray-white matter contrast. The *red contour* corresponds to the transformed gray-white matter boundary of the tissue block photograph. Jacobian determinants are in a plausible range, and the Jacobian maps are qualitatively similar. 55
3.11 Stage-1 results on a section of the orbitofrontal cortex tincture-stained with H&E, and immuno-stained for CD68 and pTDP-43. These stains exhibit a dim contrast between the gray and the white matter. The *red contour* corresponds to the transformed gray-white matter boundary of the tissue block photograph. Jacobian determinants are in a plausible range, and the Jacobian maps are qualitatively similar. 56
3.12 Comparison of stage-1 registration results with registration by ANTs. In the **top row** the results are pooled from all 14 callosal and 14 hippocampal slides. Cases where the affine initialisation failed were excluded from the ANTs results. In the **bottom two rows** are visual examples of representative registrations with ANTs SyN CC and TIRL MIND. 59

4.1 Differences in the lighting necessitates different parameters for the segmentation of the brain slices. 64
4.2 Automated initialisation and deformable registration of tissue blocks to the corresponding coronal brain slice photographs (see main text). 65
4.3 Diagram of the TIRL transformation chain that is optimised by stage 2 to align the tissue block photograph to the corresponding brain slice photograph. 68

4.4 Representative example of a successful site finding process (7 sites). **(A)** Insertion sites (*red crosses*) overlaid on the most intact brain slice photograph. The sites are supposed to be close to the centre of the corresponding tissue block. **(B)** Least intact brain slice photograph, shown for comparison. 71

4.5 Examples where manual intervention was necessary to correctly identify all insertion sites for block initialisation. Each row shows a different type of problem (these are described in the main text). *Columns 1-2*: cut-out brain slice photographs that were directly responsible for the failure. *Column 3*: erroneous result (A,B,D: the incorrect site is highlighted by the *yellow ellipses*, C: erroneous orientation of one slice (*purple*) on another (*yellow*)). *Column 4*: most intact slice with the identified insertion sites after the correction. 72

4.6 **Top:** sampling regions (A-J) with the respective tissue blocks (a-j). Note that E and G are false insertion sites with no corresponding block image. **Middle:** NMI values for all combinations of blocks (*categorical axis*) and sites (*coloured bars*). The capital letters above the chart show the region with the highest value of MI. *Red highlights* indicate where this leads to an incorrect matching. **Bottom:** The same matching process with the MIND metric. As opposed to NMI, smaller values of MIND indicate a better fit. 76

4.7 Example stage-2 registration results. **(A)** right internal capsule, **(B)** left orbitofrontal cortex, **(C)** Broca's homologue area, **(D)** left calcarine fissure. 78

4.8 Error of block initialisation relative to the final position of the block, as measured on $N = 87$ blocks. The dashed lines indicate the median (0.49 mm), the 75th (0.96 mm), and the 95th percentile (2.54 mm) distances. 79

4.9 Registration summary after 100 random initialisations in the vicinity of the automatic initialisation points. The *purple cross* indicates the true centre of each block. *Green dots* indicate successful registrations, *red dots* indicate failures. The letters correspond to the block designations in the previous experiment. The charts represent success rates as a function of the initialisation error (see main text). The gray dashed lines represent the median and the 95th percentile automated initialisation errors. 81

5.1 The difficulty of slice-to-volume matching is illustrated by comparing different cross-sections of the MRI volume. Despite the sections being relatively close, the cross-sectional anatomy varies largely between them. **(A)** coronal section, **(B)** coronal section shifted anterior by 5 mm, **(C)** coronal section rotated by 10° about the LR axis (x), **(D)** further −10° rotation about the SI axis (z), **(E)** double-rotated slice with through-plane curvature. **(F)** shows (E) with Gaussian noise (SNR = 10) to imitate low gray-white contrast. *Yellow gridlines* are provided for comparison. (Resolution: 0.85 mm/vx) 85

5.2 Diagram of the TIRL transformation chain that is optimised by stage 3 to align the brain slice photograph with volumetric MRI data. C and M refer to the brain slice photograph and the MRI volume as image functions, respectively. 88

5.3 Schematic illustration of the brain slice initialisation sequence in physical MRI space, tailored for coronal slice orientation. Alternative initialisation sequences can be defined by changing the initial parameters of the stage-3 transformations. S: superior, I: inferior, L: left, R: right. **(A)** Brain slice in pixel space (D_0), **(B)** −90° in-plane rotation after embedding at $z = 0$, **(C)** 90° rotation into the coronal plane, **(D)** Initialised brain slice in S_4 at the origin of MRI space. The position and orientation (see main text). The *lighter blue* colour indicates that the slice is now shown in radiological orientation. (Note that "surfaces" are degenerate for 2-D image data.) 91

5.4 Substage overview of the slice-to-volume registration performed by stage 3. 92

5.5 Slice-to-volume registration experiment with four series of simulated slices, corresponding to gradually increasing registration difficulty. The slices are labelled *A*–*J* from posterior to anterior. The *left column* shows the positioning of the slices in each series (superior view). The *middle column* shows an example slice (indicated by the *orange arrow*). The *yellow arrowheads* highlight the largest differences of the anatomy relative to the straight coronal slice, which therefore act as the driving force after the initialisation. The *right column* shows the evolution of the mean registration error through the 8 substages. The MRI voxel size was 0.85 mm. 98

5.6 Slice-to-volume registration experiment on simulated slices with (x:10°, y:−10°) rotations and additional quadratic through-plane curvature. Registrations are performed with both MIND and NMI used as the similarity metric. The registration accuracy is reported after each substage for both metrics. The experiment is repeated on slices with additional Gaussian noise ($SNR = 10$). MIND outperforms NMI in terms of both robustness and accuracy. 103

5.7 Evaluation of the registration accuracy on oblique simulated slices with additional through-plane curvature using different control point placement schemes. 108

5.8 Spatial distribution of registration errors (in mm) on an example oblique curvilinear slice after the 4th registration substage with *Halton* and k-means CPs. The 3-D reconstructions show the alignment of the registered slice relative to the ground truth. In both cases, the curvature is only estimated correctly under the coverage of the CPs. The difference of the curvatures is slightly larger with the k-means points. 110

5.9 Evaluation of the registration accuracy on oblique simulated slices with additional through-plane curvature using different number of control points ($4 \leq N \leq 96$). 111

5.10 The runtime of the non-linear substages 3 and 4 exhibits a quadratic dependence on the number of CPs. 113

5.11 Slice-to-volume registration result of an actual coronal brain slice photograph. The *red contour* was hand-drawn on the brain slice photograph and is provided to assess the registration accuracy. The *yellow arrowheads* indicate regions in the resampled MRI slice where the anatomy is most different from that in the tissue block photograph. Through-plane deformations prove to be essential for this slice. The corresponding Jacobian and through-plane deformation maps are shown in the *bottom row*. See main text for further details. 116

5.12 Evaluation of slice-to-volume registration errors on a real brain slice photograph, using anatomical landmarks. *Yellow dots* indicate landmarks within 1 mm from the registered slice, *red crosses* indicate landmarks further away. (**a**, **b**, **c**) compares landmark-based and automated registration in the regions with the largest error. Automated slice-to-volume registration is clearly superior in these regions, indicating relatively large errors in the annotation. 117

6.1 Three-dimensional model of the post-mortem brain showing a subset of the histological sampling sites, as they were reconstructed from the registration results. Each site is represented by the curvilinear domain of the corresponding histology image, and is therefore larger than the actual tissue section. The left-hand side of the brain was removed for better visualisation, some of the sampling sites are also not shown for greater clarity. Note that certain curves are accented because the surfaces were reconstructed from voxel-wise labels in MRI space (voxel size: 0.85 mm). 121

6.2 Overview of the combined end-to-end histology-MRI registration protocol using two photographic intermediates: the tissue block and the brain slice photographs. Identical frame colours indicate identical image domains with pixelwise correspondence. Stages 1-3 are run in parallel, and the optimised transformation chains from each are combined to transform the histology domain into MRI space, and resample the volumetric MR data onto it. The 2-D MR image is shown here as an overlay on the histolgy image with 50% opacity. . 123

6.3 Stage-1 registration results. **(A)** Representative stage-1 registration result shown by the aligned histology (*left*) and tissue block (*right*) images of a corpus callosum specimen. Grid spacing: 5 mm. **(B)** Misregistration by stretching the histology section over the side surface of the block. Confining the mask to the cut surface leads to accurate registration. **(C)** Incorrectly specified image resolutions lead to erroneous scaling. Changing the resolution input fixes the problem. The asterisk highlights a slight anatomical discrepancy between the images. 130

6.4 Stage-2 registration results. **(A-B)** Representative stage-2 registration result, shown by the alignment between the tissue block and the corresponding portion of the slice photograph for a visual cortical specimen. Grid spacing: 5 mm. **(a,b,c)** Fused red and green images showing close correspondence between perforating vessels (*arrowheads*) in the registered photographs. **(C-E)** Deflections of the cingulate gyri (*large arrow*) are responsible for the largest observed misalignments in stage 2 (*arrowhead*). **(F-H)** Misidentification of the insertion site for a hippocampal block (G), where the location of the insertion site was offset from its true location by 14.7 mm. Manual specification lead to identical high-quality registration (H). **(I-K)** Misidentification of the insertion site for a visual cortical specimen, where confounding anatomical patterns led to the incorrect estimation of the block rotation (J). Refining the rotation search led to a high-quality registration (K). 132

6.5 Stage-3 registration results. All images are shown in neurological orientation. **(A)** Tissue block photograph with highlighted gray-white matter contour. The *purple arrowheads* point at perforating vessels to assess the accuracy of the registration of the slice with the MRI volume. The *green arrowhead* shows the artificial opening of the interhemispheric fissure that leads to misregistration. **(B)** Registered 2-D MRI slice obtained with whole-slice masking. The gray-white matter contour is overlaid from the slice photograph, indicating highly accurate registration in the left hemisphere, and substantial misalignments towards the bottom of the right hemisphere. **(C)** 2-D registered MRI slice obtained with left hemisphere masking. **(D)** 2-D registered MRI slice obtained with right hemisphere masking. Both unilaterally-optimised registrations are highly accurate, as indicated by the alignment of the contours and perforating vessels. **(E-F)** The 3-D rendering of the unilaterally registered slices reveals the large difference in the slicing position and orientation between the left and the right hemispheres. 134

6.6 Histology-to-MRI registration results for various anatomical regions in the sample brain. Grid spacing: 10 mm for the thalamus, 5 mm otherwise. These results were obtained with a unilaterally optimised stage-3 registration, except for the thalamus section, where stage 3 was optimised bilaterally. The histology sections were immunostained for PLP, except for the thalamus that was inmmunostained for Ferritin. The MR images were resampled from the 0.25 mm structural (TRUFI) volume. 136

6.7 A visual comparison of the histology-to-MRI registrations on the 4 central blocks of the sample brain obtained with left-optimised and right-optimised stage-3 registration. Grid spacing: 5 mm. The histology sections were immunostained for PLP, and the MR images were resampled from the 0.25 mm structural (TRUFI) volume. . . . 137

6.8 Representative histology-to-MRI registrations for each anatomical region in subject NP004-15. Age: 65 years, Sex: male, Diagnosis: ALS. 139

6.9 Representative histology-to-MRI registrations for each anatomical region in subject NP021-15. Age: 65 years, Sex: male, Diagnosis: uncertain. 140

6.10 Representative histology-to-MRI registrations for each anatomical region in subject NP029-16. Age: 79 years, Sex: female, Diagnosis: ALS. 141

6.11 Representative histology-to-MRI registrations for each anatomical region in subject NP040-16. Age: 61 years, Sex: male, Diagnosis: no neuropathology. 142
6.12 Representative histology-to-MRI registrations for each anatomical region in subject NP096-15. Age: 71 years, Sex: male, Diagnosis: ALS. 143
6.13 Representative histology-to-MRI registrations for each anatomical region in subject NP111-14. Age: 27 years, Sex: male, Diagnosis: ALS. 144
6.14 Representative histology-to-MRI registrations for each anatomical region in subject NP135-14. Age: 76 years, Sex: male, Diagnosis: ALS. 145
6.15 (A) The source of anatomical inconsistencies between consecutive histology sections and the tissue block photograph (see the main text). (B) Application of the optional 4th stage to estimate the slicing depth and slicing angle. The colour map reveals significant pixelwise distances between the adjusted section and the histology section as it was inserted by stages 1-3. Grid spacing: 5 mm. 147
6.16 Relating histology to multi-modal MRI data. 150
6.17 Sampling an M1S1 block from the right hand area. 151
6.18 Stage-3 registration of coronal slices affected by M1S1 sampling. **(A)** Coronal slice photograph with tissue void due to M1S1 sampling. The artificial edges (*red dots*) are marked manually and correspond to the lateral surface of the hand M1S1 sampling site. **(B)** Slice mask used for stage-3 registration. **(C)** Erroneous stage-3 registration result in the absence of masking. **(D)** Accurate stage-3 registration result using the binary mask in (B). The *red contour* is an overlay of the gray-white matter boundary on the brain slice photograph. . . . 152
6.19 Different 3-D views on the manual tissue block initialisation process. The red dots were annotated in the pixel space of two neighbouring coronal slice photographs, and transformed to MRI space together with the slices. The bottom-right image shows a plane that was fitted to the transformed points to illustrate the the non-standard orientation of M1S1 block sampling. 154
6.20 M1S1 registration result. The tissue section was taken from the hand area of the left motor cortex and immunostained for PLP. The 2-D MR images were resampled at the 8 µm/px resolution of the histology image from the structural (TRUFI) volume. Post-hoc optimisation (stage 4) was applied to minimise the anatomical discrepancy in the postcentral gyrus (left). Grid spacing: 5 mm. 155
6.21 Small perturbations of the registered image during post-hoc optimisation (stage 4) lead to large changes in the appearance of the region, illustrating the difficulty of the registration. 156

6.22 A 3-D ROI was created from the Brodmann Area 4 label from the Juelich Atlas. **(A)** 3-D manually defined ROI visualised in FSLeyes. **(B)** Alignment between the registered histology image and the manually-created 3-D ROI. The alignment is close, but there is no overlap in the region of the motor cortex, only the white matter. The extent of the ROI is much larger than the histology section. . . 158

7.1 **(A)** Whole-brain histology section of the frontal lobe registered to MRI. The grid spacing is 10 mm. **(B)** Half-brain section close to the centre of the brain and the registered, resampled MR slice. The last image shows the histology image as an intensity overlay to visualise the precision of the alignment. Both histology images were prepared by the Gallyas silver impregnation technique to visualise myelinated fibres. 163

7.2 BigMac PLI. A: Transmittance image, B: Registered MRI slice after in-plane deformations. Grid spacing: 2.5 mm C: Inclination map with rotations that were introduced by TIRL: these need to be taken into account when comparing in-plane fibre orientations with diffusion maps. D: Local rotations shown in degrees. 164

7.3 Integration of multi-modal quantitative in-vivo MRI (ivMRI), post-mortem MRI (pmMRI) and semi-quantitative histology data in a single subject (see details in the main text). 166

A.1 Schematic illustration of the registration process in TIRL with the major library components involved (see main text for a description of each). 170

List of Tables

3.1 Space and image domain definitions related to the stage-1 transformation chain. 38

4.1 Space and image domain definitions related to the stage-2 transformation chain. 68

5.1 Space and image domain definitions related to the stage-3 transformation chain. 88

5.2 Comparison of the stage-3 registration accuracy on four series of simulated slices. The values of the table represent the mean registration error (in mm) averaged for all 10 slices after each registration substage. (Slice-specific MREs can be seen in Figure 5.5. These were averaged here to identify trends in the MRE, that are independent from the slice geometry.) The MRI voxel size was 0.85 mm. The highlight shows the results from the first and the repeated free-form registration substage for comparison. 99

5.3 Comparison of the stage-3 registration accuracy using MIND and NMI on simulated slices with and without Gaussian noise. The values of the table represent the mean registration error (in mm) averaged across all 10 slices after each substage of the registration. The MRI voxel size was 0.85 mm. The two free-form deformation stages are highlighted for comparison. 102

5.4 Comparison of stage-3 registration accuracy using different control point placement schemes. The values of the table represent the mean registration error (in mm) averaged across all 10 slices after each substage (0-8). The results of the two free-form deformation substages are highlighted for comparison. 107

5.5 Comparison of stage-3 registration accuracy using different numbers of control points. The values of the table represent the mean registration error (in mm) averaged across all 10 slices after each substage (0-8). The results of the two free-form deformation substages are highlighted for comparison. 112

5.6 Average runtimes of the substages in seconds as a function of control points. 113

List of Abbreviations

1-D, 2-D, 3-D, n-D	One-, two-, three- and n-dimensional
ALS	amyotrophic lateral sclerosis
API	application programming interface
CP	control point
DOF	degree(s) of freedom
FMRIB	The Oxford Centre for Functional Magnetic Resonance Imaging of the Brain
FSL	FMRIB Software Library
MCD	mean contour distance
MND	motor neuron disease
MPRAGE	magnetisation-prepared rapid gradient echo
MRE	mean registration error (averaged over all voxels)
MRI	magnetic resonance imaging
ROI	region of interest
SNR	signal-to-noise ratio
TIRL	tensor image registration library
TPS	thin-plate spline
TRUFI	true fast imaging with steady-state free precession
WSI	whole-slide image (digitised high-resolution histology image)
YIQ	the color space used by the NTSC color TV system

1 Introduction

1.1 Amyotrophic Lateral Sclerosis

Amyotrophic lateral sclerosis (ALS) [1], or motor neuron disease (MND) [2], is a rapidly progressive and fatal neurodegenerative disorder with 2-4 years median survival time after diagnosis. It normally presents at an advanced age (50-65 years) with isolated persistent muscle fatigue on one limb or the face area, which then develops into full-body paralysis, gradually interfering with vital functions, including all voluntary movements, swallowing, and breathing. To date no therapeutic approach has shown significant long-term disease-modulating effect, so the quest continues to better understand the etiology of the disease and the underlying subcellular mechanisms that could be exploited as novel therapeutic targets. ALS is a rare disease that affects approximately 1 out of 100,000 people, which is significantly less than other forms of neurodegeneration such as Alzheimer's (approximately 10,000 out of 100,000 over 65) [3] or Parkinson's (approximately 1400 out of 100,000 over 65) [4]. Nevertheless, its mysterious etiology (the unexplained 10% heritability and the higher vulnerability of certain subpopulations), its phenotypic plurality, and uncertain prognosis have kept the scientific curiosity alive around ALS since its original characterisation by Charcot in 1869 [5] [6]. A long-held view was that ALS selectively affects the primary motor neurons of the spinal

cord and the primary motor cortex in the brain. This was challenged by recent neuropathological studies [7] [8] that indicated the presence of abnormal protein substrates in various regions of the brain, suggesting a much broader extent of the disease. These post-hoc neuropathological observation and the co-occurrence of ALS with frontotemporal dementia (FTD) in affected families have led to the discovery of the ALS-FTD disease spectrum [9].

1.2 Clinical Imaging in ALS

Currently there is no definitive test to verify or exclude the presence of ALS. As a result, the diagnosis relies on clinical evaluation and a typically 1-year long observation of the symptoms, during which other causes (e.g. tumours and other forms of neurodegeneration) must be excluded. The current role of clinical imaging is to support this exclusionary process, but in the last two decades magnetic resonance imaging (MRI) has attracted considerable research interest to detect non-invasive biomarkers [10] of ALS, and the Neuroimaging Society for ALS (NiSALS) was established to share and coordinate international research in this field [11]. A recent review of ALS-related imaging biomarker candidates was published by Mazón et al [12] and earlier by Turner et al [13]. Despite promising candidates, to date no imaging biomarker has been incorporated into clinical protocols for their generally low predictive value at the individual level [14]. Research now continues to extract a multitude of more subtle disease indicators from large MRI datasets using learning algorithms [15]. The hope is that together these could provide a more sensitive and specific signature of ALS even at the individual level.

1.3 Post-mortem Imaging in ALS

Despite the relative abundance of clinical imaging studies on ALS, only very few of them have been linked to post-mortem validation. “[...] post-mortem studies often infer sequential propagation patterns without robust ante-mortem data. [...] Novel study designs, combining imaging and histological assessments, are indispensable

for the interpretation of the most commonly used imaging metrics in ALS." [16] *Histology* [17] is a gold-standard post-mortem imaging method by which small thin sections of the tissue are stained with various combination of chemicals to visualise the cellular and subcellular structure of the tissue. *Immunohistochemistry* is a specific branch in histology that employs bio-derived antibodies to selectively stain certain molecules (most often proteins) on the surface of and within cells. Traditionally these stained sections are interpreted semi-quantitatively by a trained neuropathologist to confirm or dismiss the diagnosis of MND, to evaluate the extent of the disease and to determine the disease stage (e.g. according to the Brettschneider staging [7]) by looking at samples from various anatomical locations. In recent years, stained sections have been increasingly digitised using slide scanners, allowing fully and semi-automated quantifications on each section by both commercial and open-source software packages [18].

In addition to histology, MRI in the post-mortem setting has several advantages: (1) unparalleled image resolution, (2) the use of ultra-high magnetic field strengths, (3) visualising the end stage of the disease complementing earlier *in vivo* observations. In addition to the general practice of tissue block imaging in small-bore scanners, our group has recently demonstrated the acquisition of high-quality relaxometry [19] and diffusion data from whole human post-mortem brains in a 7T clinical scanner [20] [21].

1.4 Combining Histology and MRI Data in ALS

Notable examples where ALS was studied by both MRI and histology methods include the works of Maedowcroft et al [22], Langkammer et al [23], and Kwan et al [24]. What these works have in common is that the alignment between the extracted histology section and the respective region of the MRI volume was determined manually. The manual approach has multiple shortcomings including operator bias, inaccuracies of the alignment, and most importantly a limited throughput. Consequently, the analysis could only be carried out on a limited number of samples, which probably does not represent all phenotypes of ALS. Furthermore, the MRI and

the histology data could only be compared either qualitatively or within manually defined regions of interest (ROI), which may be off-register due to inaccuracies in the alignment of the images and the segmentation.

The natural next step forward seems to be joint voxel-wise analyses between MRI and histology data in ALS, with the two modalities accurately registered. Ultimately this would not only enable us to validate MRI findings, but also to learn what data is shared between various histological and MRI representations of the same tissue. Mapping these relationships between histology and MRI could shine light on tissue microstructural changes in neurodegeneration that are responsible for MRI signal changes. The reverse mapping could supply tissue-level information based on the observed MRI signal, giving grounds to novel histology-informed imaging biomarkers with improved sensitivity and specificity for neurodegenerative conditions, including ALS.

1.5 Organisation of the thesis

This thesis aims to provide a suitable registration platform that is necessary to achieve the aforementioned goals. The pipeline should be largely automated to allow comparisons on a large number of samples, and be adaptable to the kind of images that arise from the post-mortem setting. In line with the aforementioned goals, in a recent work from our group, Pallebage-Gamarallage et al [21] outlined standard protocols for the MRI scanning, dissection, and histological sampling of whole human brains, which were optimised to study the involvement of various anatomical regions in MND. This thesis summarises the concepts and experiments related to the development of the MRI-histology registration pipeline and reports the result from its application on that dataset. The remaining of the thesis is organised into the following chapters.

1. **Chapter 2** introduces the theoretical concepts behind medical image registration, which is the broader scope that this work fits into. At the end of the

chapter, a literature review is provided that concerns the narrower context of this work: histology-to-MRI registration.

2. **Chapter 3** starts with a description of the data that was acquired as part of the post-mortem human ALS study, provides a birds' eye view of the proposed three-stage registration pipeline, and describes the development of the first stage, showing results from related experiments.

3. **Chapter 4** describes the development of the second stage and shows the results from related experiments.

4. **Chapter 5** describes the development of the third stage and shows the results form related experiments.

5. **Chapter 6** describes the combination of the three stages and the application of the pipeline to existing parts of the dataset. Includes a description of a two-stage semi-automatic version of the pipeline for registration with a single photographic intermediary.

6. **Chapter 7** reiterates the contributions of the thesis to the field, and provides future directions for research.

7. In the **Appendix** there is a description of the newly developed open-source Tensor Image Registration Library (TIRL), which provides the software backend for the proposed MRI-histology registration pipeline.

2 Medical Image Registration

2.1 Definition

2.1.1 Scope

Image registration [25] refers to a computational process whereby information from multiple (normally two) image sources is represented in the same coordinate system to establish spatial correspondence between the contents of the images. Ideally, registration is *accurate* (corresponding points in the objects, such as anatomical locations, within the images are brought into perfect alignment), *robust* (small changes in the input images make little difference in the alignment), as well as *generalisable*, such that it achieves similar accuracy and robustness on other datasets.

The emergence of modern clinical imaging techniques (computed tomography, magnetic resonance imaging, ultrasound, endoscopy, *etc.*) and the simultaneous expansion in computational capabilities led to the development of several *medical image registration methods* [26] [27] that are fine-tuned for specific types of images of the human (or animal) body. Computer-assisted diagnostics (MRI motion correction, positron emission tomography), therapy (targeted radiotherapy) and intervention (angiography, neuronavigation) all strongly rely on these techniques. In medical imaging research, and more specifically in neuroimaging [28], registration methods

are most commonly applied in motion correction, distortion correction, template creation for statistical inference testing, atlas building, and segmentation.

2.1.2 Formulation as a mathematical optimisation problem

Let the continuous $\mathcal{A} : {\Omega_a}^{p_a} \rightarrow \mathbb{R}^{l_a}$ and $\mathcal{B} : {\Omega_b}^{p_b} \rightarrow \mathbb{R}^{l_b}$ functions represent two *images*, where $\Omega_{a,b} \subset \mathbb{R}$ denote the finite physical extent of the images, and p and l are the number of physical and logical image dimensions (these are often equal between $\mathcal{A}$ and $\mathcal{B}$). Given a *dissimilarity metric* $M(\mathcal{A}, \mathcal{B})$ that quantifies the dissimilarity of two images, and a *transformation model* $T : \Omega_a^{p_a} \rightarrow \Omega_b^{p_b}$ that maps the coordinates of one image onto the other's domain, the aim of the registration process is to find the $\hat{T}$ that minimises the following *cost function*:

$$\hat{T} = \arg\min_{T} M(\mathcal{A}, \mathcal{B} \circ T) + R(T) \tag{2.1}$$

The $R(T)$ *regularisation term* can be used to impose constraints on the solution space by penalising certain types of transformations (e.g. those that are physically implausible due to the elasticity of the object, or those that bend a flat object excessively).

Unless both images are transformed, the registration has a perceived direction, such that one image remains unchanged (fixed or target image), and the other (moving or source image) is warped into the domain of the fixed image. The direction of the coordinate mapping (T) is opposite to the perceived direction of the registration, such that if "$\mathcal{B}$ is said to be registered to $\mathcal{A}$", the T coordinate mapping is from the domain of $\mathcal{A}$ onto the domain of $\mathcal{B}$.

The choice of the *optimisation algorithm* largely depends on the complexity of the transformation model. Gradient-free and gradient-based methods are both widely used in the field, as well as discrete methods in some recent applications.

The optimisation in 2.1 relies on a set of components (italicised in the above text), which are common to all techniques, hence they provide the basis for the taxonomy of medical image registration methods.

2.2 Types of input images

Depending on whether the images correspond to a single subject or different subjects, a registration may be *intra-subject* or *inter-subject* type. The latter is technically more challenging and usually less accurate, as it optimises more complex transformation models to compensate for anatomical variations between subjects.

The registration is *unimodal* if the same acquisition technique is used to collect both images (e.g. T1-weighted images at different time points). In a *multimodal* setting (e.g. functional to structural MRI registration), the correspondence between image features may be impaired due to contrast differences, requiring increasingly more complex methods to quantify the dissimilarity of the images.

Several methods specialise on input images that are of a certain dimension, and may not be compatible with other dimensions. Examples include 2-dimensional (e.g. aligning CT or MRI slices), 3-dimensional (e.g. aligning CT or MRI volumes), slice-to-volume type (e.g. aligning ultrasound to MRI), or 4-dimensional (e.g. motion correction on functional MRI time series) registration.

2.3 Transformation models

Transformations are parametric functions that map the coordinates of the target image to those of the source image. More independent parameters (*degrees of freedom*, DOF) provide greater flexibility to estimate the geometrical differences between the images, albeit at the expense of higher computational costs. A registration process may optimise more than one transformation model in order of their complexity. Transformations can be classified into *linear* and *non-linear* groups.

2.3.1 Linear

Linear transformations have relatively few degrees of freedom. Their effect is global (i.e. all coordinates are subject to the same transformation), therefore the structure of the transformed object is preserved. Translation by a vector, rotation about an axis or a point, anisotropic scaling, and shearing along an axis

are all linear coordinate transformations in Euclidean space that can be expressed uniformly by linear operators.

Combinations of translations and rotations correspond to a *rigid transformation*. With 6 DOF, a rigid transformation is sufficiently complex in most intra-subject scenarios. However, it is often used as a pre-registration step to recover large displacements before more complex transformation models are optimised. Using homogeneous coordinates, rigid transformations can be expressed using matrix multiplication:

$$R = R_z(\alpha)R_y(\beta)R_x(\gamma) = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{bmatrix} \tag{2.2}$$

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} & & & t_x \\ & R & & t_y \\ & & & t_z \\ 0 & & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = T_{\text{rigid}}\mathbf{x} \tag{2.3}$$

A 3-D *affine transformation* has 12 DOF, corresponding to the elements of a 3×4 matrix, that operates on homogeneous 3-D coordinates. Any affine matrix may also be decomposed (non-uniquely) into 3 translations, 3 rotations, 3 scales, and 3 shears, and parametrised by the respective parameters of each component (see Equation 2.4). The order of the components must be disclosed to avoid the ambiguity of the transformation, for example:

$$T_{\text{affine}} = T_{\text{shear}} \cdot T_{\text{scale}} \cdot T_{\text{rigid}} \tag{2.4}$$

If the determinant of the affine matrix is negative, the object will be flipped. The NIfTI image standard [29] stores the matrix mapping between voxel and mm coordinate systems (i.e. data and image space), where the mm coordinate system is right-handed, and hence the handedness of the voxel coordinate system is determined by the sign of the determinant (i.e. whether there is a flip or not). Accordingly, in the registration of 3-D objects, transformations can be carried out in a standard orientation (e.g. in FSL [30]), and sign flipping (between physical,

mm coordinate spaces) is normally restricted to avoid unrealistic left-right reversal. However, most 2-D image formats store no information on either the physical or pixel orientation of the image, hence this must be carefully accounted for when importing a 2-D image to be registered to a 3-D object.

All of the above linear transformations can be undone by multiplying the transformed coordinates by the inverse matrices, which can be calculated very efficiently.

2.3.2 Non-linear

Non-linear transformations (also called *deformations*) normally have high DOF, and they affect the coordinates in a positionally-dependent way, potentially disrupting the structure of the transformed object. The most general representation is a 3-D continuous displacement field, that can be parametrised in different ways. The most straightforward is by discretisation at the voxel level, when a displacement vector is defined for each pixel/voxel of the transformed object. This results in great flexibility, and no additional operations need to be carried out on the parameters to apply the transformation, but there are a large number of parameters to optimise. Alternatively, displacement vectors can be defined at a much sparser set of control points and retrieved between control points by interpolation. Several methods have been proposed that differ in the placement of the control points and the interpolation technique. Common examples in literature include interpolation by radial basis functions (ϕ, such as thin-plate splines [31]) on scattered control points (N):

$$T_{\text{rbf}}(\mathbf{x}) = \sum_{i=1}^{N} \lambda_i \phi \left(\|\mathbf{x} - \mathbf{x}_i\|_2 \right) \tag{2.5}$$

and free-from deformation [32], which is interpolation by a tensor product of trivariate B-splines on a regular grid of control points.

Alternatively, deformations can be represented as a sum of orthonormal basis functions, such as Fourier components [33], wavelets [34], where each component describes deformations at a specific spatial frequency. Kim et al [35] used second-order polynomials, however this technique has been largely abandoned in favour of the other basis functions.

Finally, deformations may also be represented as time-continuous velocity fields [36], that are integrated over time to obtain the displacement at each point. This representation is used by ANTs [37]. Although it has a substantially larger memory footprint and it involves more computations than other methods, this formalism has advantages in the estimation of large local displacements and obtaining half-way transformations.

Non-linear transformations are called *diffeomorphic* if they are continuously differentiable, invertible and the inverse is also continuously differentiable [38]. In practice, the diffeomorphic property guarantees a one-to-one and onto mapping between the image spaces, without fold-over, ensuring that no structures disappear after the transformation is applied. When fields are discretised, the diffeomorphic property holds if the Jacobian determinant of the field is greater than zero at every point. Registration tools of the FMRIB Software Library (FSL) always project fields that are not discrete-diffeomorphic to the "nearest" diffeomorphic transformation. This guarantees that the field can be inverted by estimating and inverting the local affine transformations at each voxel. Consequently, the inversion of non-linear transformations is a much more demanding process than it is for linear transformations.

2.4 Similarity metrics

Similarity metrics [39] are functions that quantify the degree of alignment between the source and target images, such that higher values of the similarity metrics correspond to better alignment. The term *dissimilarity metric* is often used to indicate an inverse relationship. Metrics can be largely classified whether they are based on landmarks or voxel intensities, although there exist hybrid measures that belong to both categories.

2.4.1 Landmark-based metrics

Landmark-based metrics rely on the definition of corresponding point features, contours (in 2-D), or surfaces (in 3-D). Salient features of the objects (*intrinsic*

landmarks) are normally segmented by a skilled operator or using automated algorithms, such as SIFT [40], the Harris detector [41], HOG [42], or SURF [43]. Depending on the image acquisition mode, auxiliary items, such as pins, stripes or markers can also be used as *extrinsic landmarks* that can be segmented more easily with minimal landmark correspondence error.

Assuming that the pairwise correspondence is known *a priori*, the most intuitive landmark-based similarity measure for N points is the sum of squared distances:

$$M_{SSD} = -\sum_{i}^{N} \|\mathbf{y}_i - T(\mathbf{p}_i)\|_2^2 \tag{2.6}$$

For contours and surfaces, the pairwise correspondence of the points are often determined on the basis of their distance, and it is revised iteratively, such as in the Iterative Closest Point algorithm [44]:

$$\mathbf{y}_i = \min_{\mathbf{q}_j \in \mathbf{q}} \{\|\mathbf{q}_j - T(\mathbf{p}_j)\|_2^2\} \tag{2.7}$$

Landmark-based similarity metrics suffer from the necessity of identifying landmarks, which is often a manually intensive or error-prone process. Their application is most common where user interaction is necessary (such as correcting artefacts in histology images [45]), or where automatic segmentation is robust enough (such as the alignment of functional and structural MRI data [46]).

2.4.2 Intensity-based metrics

These metrics are readily computed from the image intensities without an additional segmentation step, allowing higher throughput for large datasets. A large comparative study of registration methods indicated their superior performance over landmark-based measures [47]. With intensity-based metrics, the estimation of image similarity can be more robust, as it depends on all voxels as opposed to a relatively small set of points.

The most intuitive measure is the *mean of squared intensity differences* (Equation 2.8), although other variants, such as the *mean of absolute intensity differences* are also abundant in the literature.

$$M_{\mathrm{MSD}} = -\frac{1}{n} \sum (\mathcal{I}_A(\mathbf{q}) - \mathcal{I}_B(T(\mathbf{p})))^2 \tag{2.8}$$

The application of these metrics is restricted to intra-modal registration, where the corresponding image features are represented by identical intensities, or their difference is solely attributed to Gaussian noise.

A more lenient assumption is that a linear relationship holds between the intensities of corresponding image features, which is often the case for inter-modal registration (e.g. X-ray and CT, or CT and MRI). The *normalised cross correlation* metric measures the global concordance of voxel intensities:

$$M_{\mathrm{NCC}} = \frac{\sum(\mathcal{I}_A(\mathbf{q}) - \mu_A)(\mathcal{I}_B(T(\mathbf{p})) - \mu_B)}{\sqrt{(\mathcal{I}_A(\mathbf{q}) - \mu_A)^2(\mathcal{I}_B(T(\mathbf{p})) - \mu_B)^2}} \tag{2.9}$$

Local variants [48] of the same metric also exist, where the NCC is evaluated independently within small (e.g. 10×10 pixel large) patches of the images.

The most general assumption is that identical features are represented by similar intensities within each image. When the images are maximally aligned, the overlapping voxels are characterised by co-occurrences of local intensity patterns. Consequently, the voxel intensities in one image are most informative of the intensities in the other image when the images are in complete alignment. This is captured by the *mutual information* similarity metric, where the shared information content is expressed using Shannon-entropies [49]:

$$M_{\mathrm{MI}}(\mathcal{I}_A, \mathcal{I}_B) = H(\mathcal{I}_A) + H(\mathcal{I}_B) - H(\mathcal{I}_A, \mathcal{I}_B) \tag{2.10}$$

Normalised mutual information is used more commonly, as it does not depend on the overlap of the field of view between the images being registered.

$$M_{\mathrm{NMI}}(\mathcal{I}_A, \mathcal{I}_B) = \frac{H(\mathcal{I}_A) + H(\mathcal{I}_B)}{H(\mathcal{I}_A, \mathcal{I}_B)} \tag{2.11}$$

2.5 Modality-independent neighbourhood descriptor

The modality-independent neighbourhood descriptor (MIND) was proposed by Heinrich et al [50] to facilitate both the linear and deformable registration of MRI and CT volumes. It is based on the assumption that "a local representation of image

structure, which can be estimated through the similarity of small image patches within one modality, is shared across modalities" [50]. This shared representation is the MIND vector that is calculated at each pixel or voxel of the images as shown by the schematic in Figure 2.1A.

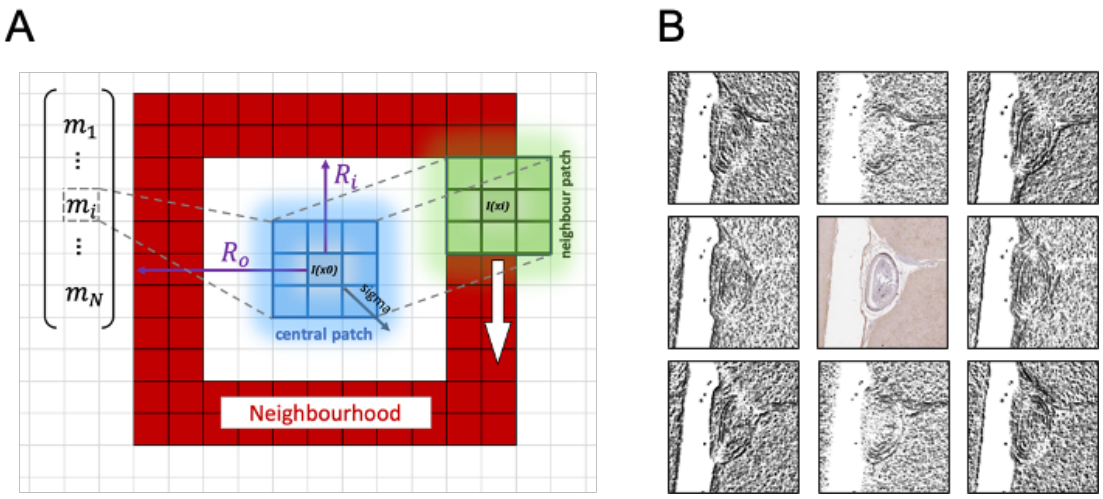

Figure 2.1: (A) Calculation of the MIND vector at $\mathbf{x_0}$ is based on the intensity values of the pixel at $\mathbf{x_0}$ and in its N-element neighbourhood (*red tiles*). In the simplest and most common case, a pixel's neighbourhood is defined by the immediate 8 neighbours in 2-D (or 26 neighbours in 3-D). Alternative definitions, excluding the diagonal elements or using different parametrisations are also possible. The schematic shows a hollow-plate neighbourhood parametrised by its inner (R_i) and outer (R_o) half-side lengths to avoid the clutter and provide greater clarity over the calculation steps. Pixel intensity relationships are calculated between image patches as opposed to individual pixels for robustness. A central patch (*blue*) of 9 pixels (or 27 voxels) is defined around $\mathbf{x_0}$, and N neighbourhood patches (*green*) of the same size are defined around the N pixels of the neighbourhood. The i-th component of the MIND vector (m_i) represents standardised *patch-wise distances* (i.e. variance-normalised sum of squared intensity differences between corresponding pixel intensities) of the central patch and the i-th neighbourhood patch. (The patch-wise distances are calculated via Gaussian convolution, hence the additional σ parameter, which effectively represents the patch size.) **(B) MIND representation of a histology image** (the image was converted to grayscale before the operation). Black-and-white images on the periphery depict the individual components of the MIND vectors (calculated for all pixels in the sample). The position of the tiles is directly related to the position of the neighbourhood patch that was used in the calculation of the respective component. Note the direction-sensitive enhancement of image edges in each tile.

Let $u, v \in \mathbb{N}$ denote the coordinates of the pixel at $\mathbf{x_0}$ with an arbitrary-shaped neighbourhood $\mathcal{N}$ (Figure 2.1A, red tiles). Let the elements of the neighbourhood

be indexed $\mathcal{N}_{1...i...N}$ and defined as follows:

$$\mathcal{N}_i(u,v) = \mathcal{I}(u + l_i, v + m_i) \tag{2.12}$$

where $l_i, m_i \in \mathbb{Z}$ represent the offset of the neighbourhood pixel from the central pixel along each image dimension.

The most commonly used neighbourhood layouts include the 8 immediate neighbours, or the 4 immediate neighbours (i.e. excluding the diagonals) in 2-D, and equivalently, the 26 or 6 immediate neighbours in 3-D. More extended shapes increase the computational demand of calculating the similarity metric and usually have diminishing returns in terms of registration accuracy. The ideal shape of the neighbourhood should be determined empirically for each application, although they tend to lead to similar results [50].

Let $\mathcal{P}_0$ and $\mathcal{P}_1 ... \mathcal{P}_N$ be identical square patches around the pixel at (u, v) and each pixel in $\mathcal{N}$. Let the pixel elements of the i-th patch be indexed as $\mathcal{P}_i^{(1...j...P)}$.

Heinrich et al defined the *patch-wise distance* between the central patch ($\mathcal{P}_0$) and an arbitrary neighbourhood patch ($\mathcal{P}_i$) as follows:

$$D_i = \sum_{j=1}^{P} \left(\mathcal{I}(\mathcal{P}_0^{(j)}) - \mathcal{I}(\mathcal{P}_i^{(j)}) \right)^2 \tag{2.13}$$

For images that have slow-varying intensity overlays (e.g. a coil-sensitivity bias in MR volumes), or have an uneven distribution of image noise (e.g. noise amplification in multi-coil MR acquisitions), the patch-wise distances may exhibit regional disturbances, which are unrelated to the "interesting" point-like features, sharp edges and gradients of the image. The components of the MIND vector ($m_{1...i...N}$) are therefore defined as *standard patch-wise distances*:

$$\text{MIND}_i(u,v) = \frac{1}{\lambda} exp\left(-\frac{D_i(u,v)}{V(u,v)} \right) \tag{2.14}$$

where the V normalisation factor is the mean of all patch-wise distances in the neighbourhood:

$$V(u,v) = \frac{1}{N} \sum_{i}^{N} D_i(u,v) \tag{2.15}$$

and λ is another normalisation factor that is equal to the largest MIND component at each pixel, so that the range of the MIND components is between 0 and 1.

The MIND representation of the source and target images is a vector of values at each location for each image, and it is the components of this vector that can then be compared separately. The registration is therefore reduced to the unimodal case (as the values of corresponding elements in the vector should be equal across images when well aligned). The simplest way to define a dissimilarity metric for MIND representations is by taking the squared Euclidean distance between corresponding MIND vectors. In the following we will refer to this as the "MIND metric". The MIND metric, which has analytical derivatives, can be used for efficient optimisation of the transformation parameters. The main advantage of the MIND metric over mutual information is that it retains the local orientation information that is shared across the images. MIND has some similarities with a feature detector, but with features that are identified at all voxels, which makes the MIND metric more of an intensity-based similarity metric.

2.6 Optimisation

At the core of any registration process is the optimisation algorithm that aims to determine the specific set of parameters for the transformation model that is associated with the smallest value of the cost function. The computational demand of the optimisation increases with the number of parameters, as the dimensionality (hence the size) of the search space becomes larger. This and other factors (e.g. difficulty of implementation) have motivated the use of different mathematical optimisation techniques in the context of image registration. The majority of these algorithms are iterative, i.e. they infer the optimum set of parameters by gradually updating an initial estimate. Different iterative methods use different strategies to compute the parameter updates, and they can be classified [39] into *gradient-based*, *gradient-free*, *discrete* and *heuristic* categories. As an alternative to iterative methods, deep-learning models have recently emerged [51]. These can calculate the optimum parameters in a single step, which makes them fast

in application, but they normally require extensive training on a large number of domain-specific images, and they are currently more effective in the field of unimodal registration [51]. Therefore, in the scope of this thesis, only well-proven gradient-based and gradient-free methods will be used.

2.6.1 Gradient-based methods

These methods [52] calculate the parameter updates at each iteration as a product of a search direction ($\mathbf{h}$) and a step size (α):

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \alpha \mathbf{h}_k \tag{2.16}$$

To calculate the search direction, all methods in this category rely on the first or higher-order derivatives of the cost function ($C(\mathbf{p})$) with respect to the transformation parameters ($\mathbf{p}$). The simplest strategy is called *steepest gradient descent*, where the search direction is equal to the negative first derivative of the cost function:

$$\mathbf{h}_{\mathrm{GD}} = -\nabla_{\mathbf{p}} C \tag{2.17}$$

The step size is adjusted in each step by line search to ensure that consecutive parameter estimates are associated with monotonically smaller values of the cost function. The descent direction at each iteration is orthogonal to the previous one, hence the algorithm approaches the optimum in a zigzag pattern, which may become inefficient after the initial few steps. In practice, the gradient is often estimated by finite difference methods, which may also impair the robustness of the method.

In *conjugate gradient methods*, the first step is taken along the negative local gradient, then all subsequent search directions are computed from the residual and the previous search directions, such that all search directions are conjugate. In certain cases this method requires fewer steps to reach or get close to the optimum than gradient descent, but it still can be a lengthy procedure.

Newton's method is based on the second-order Taylor expansion of C around $\mathbf{p}_0$. As such, the search direction is given by the solution to the following equation:

$$\mathbf{H}\mathbf{h}_N = -\nabla_{\mathbf{p}} C \tag{2.18}$$

where **H** is the Hessian matrix (second derivative of the cost function). While Newton's method achieves quadratic convergence, the exact calculation of the Hessian is often computationally expensive, and therefore limits its practical application.

Quasi-Newton methods employ different ways to estimate the Hessian matrix without having to calculate second derivatives, which makes them much faster than gradient descent and Newton's method. Commonly used variants are BFGS, Gauss-Newton and the Levenberg-Marquardt methods [53].

2.6.2 Gradient-free methods

Gradient-free methods can be applied to optimise a wider range of cost functions, including those that are noisy and non-differentiable. They also have a clear advantage over the methods of the previous group when analytical derivatives are hard to implement and the derivatives cannot be estimated numerically with sufficient precision.

The *Nelder-Mead method* initialises a simplex in the parameter space, evaluates the cost function at the vertices and marches the simplex by repositioning the worst vertex at every iteration until convergence is reached. This method is more robust against local minima than gradient descent, which makes it a popular choice, but the time required for its convergence tends to be long.

Powell's method explores the parameter space by performing bidirectional searches along a set of predefined search vectors. At every iteration, the parameter updates are determined as a combination of the search results, and basis vectors are updated using the information about the vector that contributed most to the update in order to generate a new set of search directions for the next iteration. Powell's method is "generally robust, accurate and exhibits fast convergence" [39].

The *NEWUOA* [54] and *BOBYQA* [55] *methods* (also proposed by Powell) were found to be superior to Powell's direction-set method by Osechinskiy and Kruggel [56] for the slice-to-volume registration of histology and MR data. These perform a local quadratic approximation of the cost function using a limited number of

cost function evaluations, and BOBYQA additionally allows the specification of bounds on the parameter values.

2.6.3 Discrete methods

The search space for optimal transformation parameters can be effectively reduced by discretising the parameters and testing a finite set of parameter constellations. For relatively few parameters, a global optimum can be found very efficiently by brute-force testing. For a large number of parameters, however, more sophisticated graph-theoretical approaches have been developed [57] [58], that conceptualise registration as a discrete Markov Random Field labelling problem.

Discrete methods have multiple advantages over continuous methods [39]: (1) they are inherently gradient-free, (2) they are more robust against local minima, (3) efficient parallel implementations are possible, (4) prior information on the parameters can be incorporated, as the search range and resolution of the parameters are explicitly controlled by the discretisation. Two major drawbacks of discrete methods are that they become inefficient for large search spaces, such as in 3-D registrations, and that their accuracy is strongly limited by the discretisation, thus more sophistication is necessary to reach the accuracy of continuous methods.

2.6.4 Heuristic methods

Methods in the other categories have the common limitation that they tend to converge on inferior solutions when the initialisation is far from the global optimum, and/or the cost function has several local optima. Heuristic methods, such as *simulated annealing* and *differential evolution* can be used in conjunction with other methods to find the global optimum, but these methods are generally slower, and a good initialisation or registration at coarser image resolutions with pyramidal subsampling are the preferred choices to avoid local minima.

2.7 Previous work in Histology-to-MRI registration

It is important to distinguish between two main approaches to MRI-histology registration based on how the histology data is collected, as it largely determines how the registration is performed. Over the next few paragraphs we shortly review previously proposed registration methods for (1) dense systematic histological sampling, and (2) stand-alone histological images.

2.7.1 Volume-based methods

Methods of the first kind are well-developed with numerous examples [59]–[69] from as early as 1994. A comprehensive review of the methods in this category can be found in Pichat et al [70]. For these methods, tissues must be frozen or embedded in a rigid medium and sectioned at regular intervals. Most commonly the tissue block face is photographed after each section to serve as a rigid reference. Distortions of the thin tissue sections are compensated by 2-D deformable registration to the corresponding rigid reference, which are subsequently stacked to create a volume of photographic/histological data. This volume is later registered to the MRI data using standard 3-D registration tools such as ABA [71] or ANTs [37]. As a novelty, Iglesias et al [72] recently demonstrated accurate (0.5 – 2 mm) 2D-to-2D histology-to-MRI registration without photographic intermediates. Assuming perfect slice correspondence, they mapped sequentially sampled whole-brain histology images to an MRI volume. Complementary to this work, Pichat et al [73] proposed a method for direct histology-to-MRI registration between small histological samples (as opposed to whole-hemisphere images) and corresponding MRI slices via automated affine-invariant shape matching, but only reported preliminary results that required manual MRI slice matching. In their most recent work [74] they present a computational pipeline that can be used with small-format histology to provide whole-brain coverage. This pipeline also depends on stacking histology images.

Perhaps the biggest advantage of the methods in this category is that the 3-D geometrical correspondence of the sections is accurately preserved. With the use

of a rigid reference, the "banana effect" [75] in which an accumulation of small shifts between adjacent slices results in a shearing in the third dimension, may also be avoided. However, given the size of the human brain, the acquisition of systematically sampled histology data requires bespoke slicing and stain automation hardware, none of which is readily available at most neuropathology facilities. This approach is therefore better suited to study the brains of small animals, infant brains, hemispheres, or, with substantial time and workforce commitment, a single human brain [76], [77]. A more recent work by Alegro et al [78] demonstrated the automation of whole-brain histology reconstruction and registration with MRI using ANTs. Most recently, Ni et al [79] proposed a parallel block-matching algorithm that is more memory-efficient than ANTs, and demonstrated successful registrations of terabyte-sized 3-D histology and MRI datasets. Both recent developments constitute an important step towards large-scale systematic multi-subject studies.

2.7.2 Single-slice methods

For stand-alone histological images of isolated samples (i.e. a single, small-format slide), a direct slice-to-volume registration must be employed. Despite the fact that almost the entire body of histology images that have ever been created in neuropathology facilities for expert interpretation exist in this format, suitable registration methods are disproportionately under-represented in the literature. A relevant survey of slice-to-volume registration methods was recently published by Ferrante et al [39]. Here we only summarise those works that have been specifically considered for the registration of MRI and histology.

There are multiple unique challenges associated with slice-to-volume registration of small histology images. Firstly, accurate localisation is crucial, as a small 2-D slice may constitute a relatively good fit at many different locations in 3-D space. Secondly, the parameters that are necessary to align a potentially distorted 2-D image in 3-D space have a vast search space, and a high propensity

to converge to local optima. Consequently, most reported pipelines are semi-automatic, requiring accurate manual slice initialisation or annotating correspondent anatomical landmarks.

Kim et al [35] reported the first relevant slice-to-volume registration approach between post-mortem brain slice photographs and ante-mortem MRI, using 2nd - and 3rd -order polynomial parametrisation for in-plane and out-of-plane deformations. The method was later used by Singh et al [80] to register histological images of vascular lesions to in-vivo MRI data using photographic intermediates, reporting an overall registration accuracy of 3-8 mm. Meyer et al [81] reported a semi-automated registration method to align a histological section of murine glioma to in-vivo MRI using high-resolution ex-vivo MRI as an intermediate reference volume. As an improvement to using polynomials to parametrise deformations, they used thin-plate splines (TPS) to correct for both in-plane and out-of-plane deformations of the histological section. However, the accuracy of their method was not mentioned, and the code was not published.

Osechinskiy et al [56] registered histological sections of whole hemispheres to 3-D MRI data and conducted a comprehensive analysis of cost functions, optimisation and transformation methods. The best results were achieved by using TPS-based parametrisation of a 3-D deformation field, optimised by the NEWUOA algorithm [54] for Pearson's correlation of the images. As a further refinement to the technique, they devised a novel similarity metric based on the correspondence of grey-white matter boundaries [82]. Unfortunately, the authors never made their implementation publicly available for reuse, prohibiting the test of novel cost functions, including the one that is based on the MIND metric [50].

The registration of small histological sections (as opposed to whole-hemisphere images) was studied by Ohnishi et al [45]. Using manual landmarks, they stitched together smaller histological images into a full hemisphere, and subsequently registered that to a photograph of the coronal section of the hemisphere, which was further registered to 3-D MRI. Neither in-plane nor out-of-plane deformations

were considered in the slice-to-volume step; the authors instead recommended using specialised hardware to avoid distortions.

2.7.3 Considerations for a new image registration framework

In both categories of problems described above, the pipelines mentioned so far were built for a specific purpose and have not generally been released in the form of open-source software. As a consequence, new experimenters are repeatedly required to invest time and effort into extensive hardware and/or software development [83], which negatively impacts large-scale validation studies. Recently, Majka et al [84] released Possum, an open-source framework for reconstructing volumetric histology data, which is a great step toward standardising this aspect of stack-based MRI-histology registration. As mentioned before, Alegro et al [78] reported an automated pipeline for serial histological sections. To the best of our knowledge, similar automated software tools for slice-to-volume registration of small stand-alone histological images to MRI data are still scarce. HistoloZee is an excellent interactive visualisation and registration tool, however its automation and slice-to-volume registration capabilities are limited and it is not currently managed as an open-source project [85].

In the present work, we aim to address this need by describing a new open-source image registration framework (Tensor Image Registration Library; TIRL) that aims to integrate previously published methods in a single package, providing a customisable workflow that is compatible with most common image formats for MRI, histology and photographs. Based on this framework, we propose a fully automated registration pipeline to align small histological sections with volumetric MRI data using photographic intermediates. As an additional novelty, we demonstrate by example that in the case of free-hand brain cutting, involuntary deflections from the slicing plane are large enough to disrupt the anatomical correspondence between histological sections and visually matched slices of the MRI volume, and explain how these distortions are compensated within the proposed pipeline.

3

Stage 1: Histology-to-block registration

Contents

3.1 Dataset 25
3.2 Histology-to-MRI Registration in three stages 27
3.3 Rationale 28
3.4 Distortions and artefacts of the input images 29
3.5 Image pre-processing 31
3.6 Registration approach 33
3.7 Evaluating the accuracy of the registration 47
3.8 Exploring the effects of regularisation 49
3.9 Registering sections with different stains 54
3.10 A comparison with ANTs SyN 58

3.1 Dataset

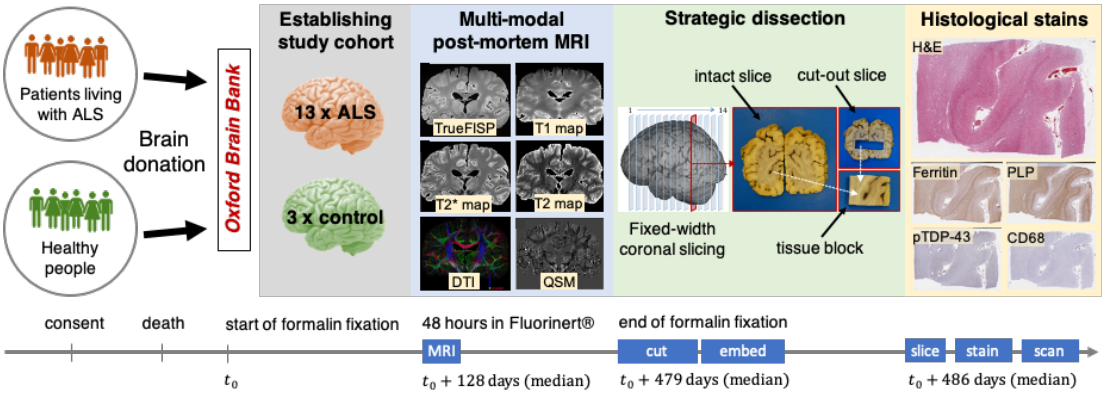

Figure 3.1: Overview of the post-mortem biomarker study data collection pipeline.

Figure 3.1 summarises the collection of the imaging data that served as a starting point for the present study. All data was collected and used according to the Oxford Brain Bank's (OBB) generic Research Ethics Committee approval (15/SC/0639). Written informed consent was obtained by the OBB from all participants of this study. Thirteen formalin-fixed post-mortem brains with neuropathologically and clinically verified MND were obtained from the cases of the OBB between 2014 and 2017. The median age of the donors was 65.5 years at death (full range: 27-77 years), and ten of them were males, two of them females. An additional three brains with no neuropathological hallmarks or clinical records of neurodegeneration were obtained as controls (age at death: 61, 76, 89 years; 2 males, 1 female). The post-mortem interval of the brains varied between 1 and 7 days (median: 3 days). The brains (denuded of the dura mater) were immersed in 10% neutral buffered formalin for at least 1 month to allow even preservation of the tissues throughout the full volume of the brain.

3.1.1 Post-Mortem Magnetic Resonance Imaging

A number of different types of images (sequences) were acquired for the post-mortem MRI. This included images that reflected the magnetic relaxation properties of the tissue (T1, T2, T2*) either in the form of weighted-intensities or quantitative reconstructions. These relaxation times are influenced by the micro-environment

of the different gross tissue types (grey matter, white matter) and are used to distinguish these tissue and any pathological changes, such as iron deposition, to them. In addition, diffusion-weighted images were acquired that are sensitive to the axonal fibre distributions, reflecting information about directions, relative sizes of fibre populations, and axonal integrity. For the purposes of this work we will largely be interested in the contrast between the gross tissue types and how this can drive registration, and be used to measure its performance.

Before scanning, each post-mortem brain was placed into a brain-shape plastic container to prevent large deformations and the container was filled with Fluorinert to suppress the background signal. Scans were performed on a clinical 7T Siemens MRI scanner at the Wellcome Centre for Integrative Neuroimaging (University of Oxford) using an optimised 48-hour acquisition protocol yielding quantitative T1 and T2 maps at 1 mm isotropic resolution, T2* and susceptibility maps at 0.5 mm isotropic resolution, DWI at 0.85 mm isotropic resolution and a bSSFP anatomical reference scan at 0.25 mm isotropic resolution (also referred to as 3D-TRUFI) [20], [21]. The MRI images were aligned and post-processed with an in-house pipeline [86] [87]. For the purposes of this work, only the anatomical reference scan was used, because it exhibited the highest contrast between grey and white matter.

3.1.2 Dissection photographs

The formalin-fixed whole brains were subsequently dissected at the Neuropathology Department of the John Radcliffe Hospital (Oxford). Following an optimised whole-brain sampling protocol [21], each hemisphere was sampled directly on its dorsolateral surface to extract 3 tissue blocks from the primary motor (M1) and the primary sensory (S1) cortices, corresponding to the somatotopic representation areas of the hand and the upper and the lower limbs (we will refer to the extracted blocks later as the “M1S1 blocks”). The brain was subsequently sliced into approximately 1 cm thick coronal sections by hand, starting from the plane of the mamillary bodies toward the anterior and posterior poles of the brain. The total number of slices (13-17) varied with the size of the brain. The middle slices were affected by the

sampling of the M1S1 blocks, presenting with bilateral damage corresponding to the 3+3 sampling sites. From predefined anatomical locations in each slice, one or more smaller tissue blocks were extracted by knife section. The size and shape of the blocks varied across sampling sites, but most of them were not larger than a few centimetres. The whole process was carefully documented by systematically capturing photographs of (1) both the anterior and posterior surfaces of the coronal slices, (2) each coronal slice after the extraction of a new tissue block, and (3) both the anterior and posterior faces of all tissue blocks (Figure 3.1). The size of the raw photographs was 5472×3648, and their resolution was later recorded from a ruler in the images as approximately 55 µm/pixel.

3.1.3 Histology

Finally, all tissue blocks were embedded in paraffin and sectioned on a microtome at 6-10 µm thickness for histology (the exact thickness depended on the staining protocol). Tissue sections were stained by various histological methods, including standard haematoxylin and eosin (H&E) and immunochemistry for proteins of interest such as ferritin, myelin proteolipid protein (PLP), activated microglia and macrophages (CD68), phosphorylated TAR-DNA binding protein-43 (pTDP-43) and pan microglia (Iba-1). Specifically, for the purpose of registration with MRI, additional LFB+PAS (Luxol fast blue combined with the periodic acid-Schiff procedure) staining was performed on two blocks of a single brain. Digital whole-slide images were created in SVS format using an Aperio ScanScope slide scanner at $\times 20$ objective magnification, yielding a typical matrix size of approximately 60000×45000 at full resolution (approximately 0.5 µm/pixel).

3.2 Histology-to-MRI Registration in three stages

Figure 3.2 outlines the plan for the registration of the histology images to the structural MR volume using the dissection photographs as intermediates. It is clear that the pipeline has to solve several compatibility issues, because of the variety of image types and formats used. It additionally has to address the two

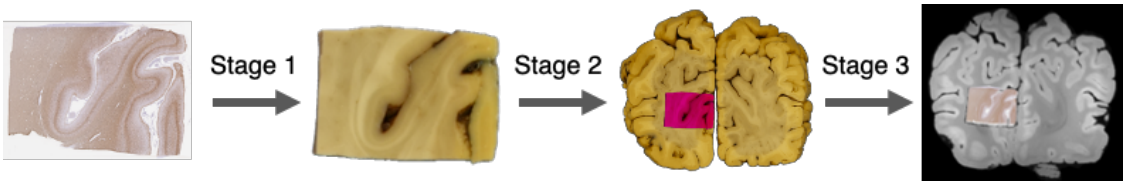

Figure 3.2: Overview of the proposed three-stage histology-MRI registration pipeline. The two photographic intermediates help with the automation of the pipeline.

orders of magnitude difference in the resolution and the dimension mismatch that is between the start and the endpoint.

A general overview of the novel image registration framework (TIRL) that was developed by the author from scratch is provided in the Appendix. All stages of the proposed pipeline were also developed by the author on top of this framework and these are covered in detail throughout the chapters of this thesis.

The rest of this chapter explains the rationale behind the first stage, describes the registration approach and reports the experiments that were carried out to verify the accuracy and the robustness of the implementation. The later parts of the pipeline, including the combination of the three stages will be covered in the subsequent research chapters.

3.3 Rationale

The purpose of stage 1 is to establish an accurate alignment between a stained and digitised histological section and the corresponding tissue block photograph. It is assumed that the photograph of a freshly cut tissue block is a less distorted representation of the anatomy than the histology image, as the latter is obtained through various processing steps, during which a thin section of tissue may deform in plane and might also tear or fold. In addition to these in-plane deformations, the total distance between a histology section and the corresponding surface in MRI also includes deformations that are perpendicular to the histology section. These arise from the bulk deformations of the brain as it is transferred from its scanning position to the cutting position. Due to bulk deformations, the cutting planes are not only different from the MRI slices, but they might be curved if the brain deforms

perpendicular to the cutting plane, hence the perpendicular component of the total deformations. Estimating the histology-related deformations is therefore significantly easier than estimating the total deformations, as they are smaller, and most non-linear registration methods tend to perform better when displacements are small.

3.4 Distortions and artefacts of the input images

Both the histology image and the tissue block photograph suffer from distortions and artefacts that are specific to each. Being knowledgeable about these is also useful to minimise them during the acquisition of the images, which can greatly help increasing the robustness and the accuracy of the stage-1 registration.

3.4.1 Histology

Typical distortions of the histology image can be attributed to specific steps in the pre-processing of the tissue. First, tissue blocks are embedded in paraffin to facilitate microtome sectioning. Paraffin embedding is preceded by the dehydration of the tissue block in alcohol/xylene bath. Dehydration leads to a substantial shrinkage of the blocks: approximately 19% by length [88]. When the paraffin-embedded block is sectioned in a microtome, slices are immediately transferred to a water bath allowing the surface tension of the water to prevent the sections from curling. Free-floating sections are then mounted on glass slides by hand. At this step, the tissue is stretched (global deformation) and parts that are more mobile, such as a gyrus might bend (local in-plane deformations). Artefacts from this step include tears, wrinkles and folds. Further artefacts may occur during the staining process, as some parts of the tissue might attain excess stain, or may remain either completely or partially stain-free (when multiple stains are used, only one of them is visible), depending on how the slides are mounted, and how long they are exposed to chemical reagents. To keep these artefacts at minimum, it is crucial that lab standards are closely adhered to and all samples in an analysis are produced in the same batch. When the slide is covered with a cover slip and sealed airtight with resin, tissue debris and bubbles might remain entrapped, contributing to the artefacts in the

digitised histology image. Artefacts may also arise when the slides are digitised. We noticed that our slide scanner employed automatic backlight compensation based on a small number of preselected regions of the slide. If any of these areas were corrupted by staining artefacts or tissue debris, the scanned image presented with a visible stripe pattern. Finally, it should be noted that slides are normally scanned with a non-white bright background, which makes image segmentation harder, as some parts of the tissue might blend into the background colour.

3.4.2 Tissue block photograph

When a photograph is taken of a tissue block, the image of a 3-D object is projected onto the sensor of the camera to produce a 2-D image. Unless the optical axis is perpendicular to the cut surface of the tissue block, the surface will appear in perspective projection, which may be more or less severe depending on the angle. As opposed to the histology and MRI images, the resolution of photographs is not fixed: it depends on the distance between the camera and the object, which may vary between images. These effects should ideally be minimised by mounting the camera and including a visual scale (e.g. a ruler).

If the tissue block has a slight pyramidal shape (i.e. its base area is larger than the top), portions of its *side surface* may be visible in the photograph that are not represented in the histology image, because these are beneath the slicing plane. More of the side surface becomes visible when the block is further away from the optical axis, therefore blocks should be positioned at the centre of the image. When multiple blocks are photographed at the same time, it is inevitable that some of them will appear with their partially visible side surface.

Residual fluid (e.g. formalin) on the block surface may produce bright spots (reflections), which may hide relevant anatomical contrast. The strength of anatomical contrast may also vary based on the lighting conditions and the colour temperature of the photograph.

Finally, it is an advantage of the photographs that an appropriate background with a uniform and distinct colour can be chosen to facilitate the segmentation of the tissue in the post-processing.

3.5 Image pre-processing

3.5.1 Tissue-block photograph

The raw-format photographs were corrected for lens distortion and converted to TIFF format by the author using the Darktable software [89]. Figure 3.3 shows an example of the original photographs that were taken of two tissue blocks. The image is approximately ∼3600 x 5400 pixels large, corresponding to a ∼50 µm/px resolution. In the majority of the cases, multiple tissue blocks were photographed at once (typically 2-4), together with auxiliary items, such as a ruler (allowing accurate determination of the image resolution), the subject label, and tissue cassettes. After the photographs were taken, the blocks were returned to their labelled cassettes and embedded in paraffin in preparation for histology. For the registration, the blocks had to be extracted from the images, and segmented from the blue background.

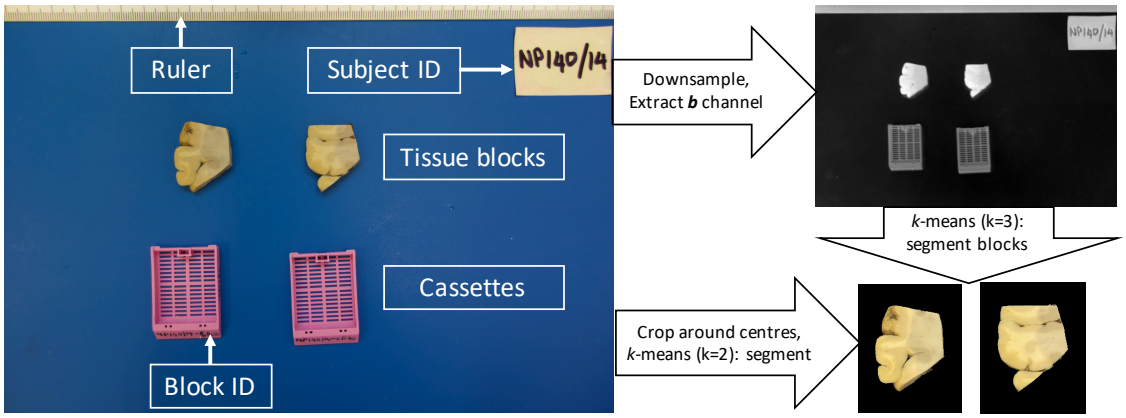

Figure 3.3: Segmenting tissue blocks from the original autopsy photographs.

To extract the tissue blocks, the images were downsampled by a factor of 16 (0.8 mm/px), and converted to the CIE-L*a*b colour space. The b channel was extracted, intensity values were converted to Student t-values, and k-means clustering ($k = 3$) was performed on the t values. Cluster probabilities were

calculated for each pixel, and an initial segmentation was performed by highest class probabilities. The cluster with the highest t-value was associated with the tissue blocks to create a binary mask for the blocks. The blocks were identified by connected component analysis on the mask, and the coordinates of their centroids were scaled back to the original resolution. To exclude any erroneously identified objects, further filtering steps were employed. A rectangle with $1/4^{\text{th}}$ the size of the image was defined around the centre of the original image, and any objects with a centroid outside of this rectangle was discarded. Furthermore, any object smaller than 1000 pixels (2.5 mm^2) was discarded as well. A bounding box was defined around the segmented blocks, and expanded in all four directions by 25%. The inflated regions were cropped from the original image, and the block mask was applied to remove the blue background in the output images.

The pre-processed block images were fed into the stage-1 registration pipeline along with information about their resolution. The latter was taken as a constant for each autopsy session and was determined manually by reading the ruler and dividing by the corresponding number of voxels along a straight line in one of the images using the ImageJ software [90].

At the start of the stage-1 registration routine, the pre-processed RGB images were reduced to grayscale by taking the luminance values (Y) of the pixels in YIQ colour space:

$$\begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.5959 & -0.2746 & -0.3213 \\ 0.2115 & -0.5227 & 0.3112 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} \tag{3.1}$$

3.5.2 Histology

Histology slides from the post-mortem study were digitised using an Aperio slide scanner to SVS format with JPEG compression at x20 magnification. This yielded whole-slide images (WSI) with a typical size of 45,000 x 60,000 pixels, corresponding to a constant resolution of 0.5 µm/px across all images. The typical size of the SVS files were between 500 MB and 2 GB, depending on the FOV. Histology

images were imported into the stage-1 registration pipeline at 8 µm/px resolution (image size ∼2800 x 3750 pixels).

As part of the stage-1 registration routine, further three pre-processing steps were applied to the histology images.

1. The images were downsampled to match the resolution of the corresponding tissue block photograph using bilinear interpolation. Before pixel values were resampled on the coarser grid, the original images were convolved with a Gaussian kernel (separately for each colour channel) as in Jenkinson [91], to ensure that all original pixels contributed to the pixels of the low-resolution image, which was eventually used in the registration. The downsampling factor was calculated as a fraction of the image resolutions: $f = 50/8 = 6.25$

2. The downsampled histology image was converted to grayscale by transforming RGB vectors to the YIQ colour space and keeping only the luminance values (Y) of the pixels, as defined above (3.1) for the tissue block photographs.

3. The downsampled grayscale histology images were padded on all four sides by $1/6^{\text{th}}$ of the respective image dimension. The padded region was filled with the maximum value of the image, in keeping with the bright background of the histology images. The padding was found necessary, as the tissue sections were often confined by the image borders. When a histology image with a tight FOV is oriented perpendicular to the tissue block photograph, the calculation of image similarity is flawed because a large non-overlapping portion of the images is ignored from the calculation, leading to erroneous registration. Padding the histology image prevents this behaviour.

3.6 Registration approach

The registration in stage 1 is formalised such that the points of the histology image (H) are non-linearly mapped onto the points of the tissue block photograph (B). The final alignment is achieved by optimising a chain of transformations

through four substages: a rotation search, plus a similarity, an affine, and a non-linear transformation. Understanding further details of the substages warrants a description of the coordinate spaces, image domains and transformations used throughout this stage.

3.6.1 Transforming the images to physical space

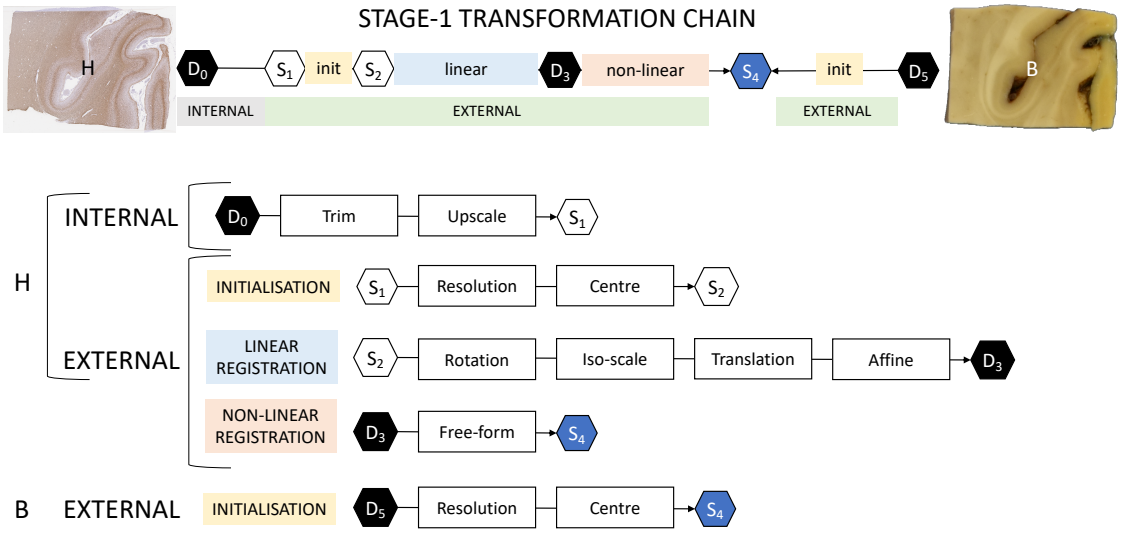

Figure 3.4: Diagram of the two TIRL transformation chains that are used in stage 1 to align the histology image (H) with the tissue block photograph (B). *Black hexagons* D_0 and D_5 represent the pixel domains of the images that the transformations are applied to. (These correspond to actual `Domain` objects in the implementation of stage 1.) D_3 is the domain of the non-linear transformation in physical space, which is instantiated from D_0 with the internal, init and linear subchains. *White hexagons* S_1 and S_2 are virtual entities, representing the intermediate, partially transformed histology domains that stand as the source and target domains of the init subchain. S_4 (*blue hexagon*) denotes the (virtual) common target of the H and B chains.

Figure 3.4 shows two chains of coordinate transformations: one that transforms the domain of the histology image ($D_0 \subset \mathbb{R}^2$) (we will refer to this as c_H), and another that transforms the domain of the tissue block photograph ($D_5 \subset \mathbb{R}^2$) (we will call this c_B). The two chains map the pixel coordinates of each image ($\mathbf{x}_H \in D_0$ and $\mathbf{x}_B \in D_5$) into a common space, which is denoted by S_4 (*blue hexagon*). The S_4 space represents the real world (which we hereafter refer to as *physical space*). In the specific case of stage 1, the physical space is a 2-D Euclidean metric space, where the image coordinates are expressed in millimetres: $S_4 \subset \mathbb{R}^2$

By mapping both images into physical space (S_4) as part of the registration process, we are accounting for three frequently encountered differences between the input images.

1. If the *number of pixels is different* between the two input images, pixel-by-pixel correspondence cannot be achieved by stacking the images. Yet it would be necessary to establish such a correspondence to calculate the similarity of the images.

2. If the input images had perfectly conformant shapes, *large differences in the image resolution* could still drive the registration to a local optimum, for instance where one whole block is aligned with a small rectangular pattern that exists in the other image.

3. Even with identical pixel shape and resolution of the input images, the regions of interests could still be in opposite corners. When such images are superimposed, the *ROIs are either completely or largely disjoint*. Instead of overlapping with each other, each ROI faces a uniform background area of the other image. In this setting, small displacements of the ROIs will have no effect on the cost, as the background has a constant intensity value (e.g. 0) in many cases. This makes cost derivatives flat, which in turn leads to a premature termination of the registration, leaving the ROIs far apart from each other.

To understand how the transformation chains c_H and c_B address these issues, we will first take a look at the bottom half of Figure 3.4 that breaks down the chains into subunits and shows elementary transformations as white rectangles. Each of these rectangles corresponds to a separate TIRL `Tranformation` object. (TIRL allows these to be optimised in any order or combination with other `Transformation` objects irrespective of their position in the chain.)

3.6.2 Deconstructing the transformation chain of the tissue block

The chain of the tissue block photograph (c_B) is an injection that maps the pixel coordinates of the tissue block photo into physical space:

$$c_B : D_5 \rightarrow S_4 : \mathbf{x}'_B = c_B(\mathbf{x}_B) \tag{3.2}$$

The total mapping is decomposed into two elementary transformations. The first is a scaling component (*Resolution*), that multiplies the pixel coordinates by the pixel size (s_x, s_y) in millimetres, which is determined from manual input (see Section 3.5). The second component is a translation (*Centre*), which moves the weighted centre of the tissue block photograph to the origin of S_4. To calculate the weighted centre, pixel coordinates of B are multiplied by the luminance of each pixel, and averaged across the whole image and scaled by the resolution. By moving the weighted centre to the origin of S_4, we fix the position of the tissue block ROI, thereby addressing the issue of having ROIs in arbitrary locations in the input images.

Both transformations are in the *external* ("user-managed") part of the chain c_B. (Note that *user* in this context refers to the user of the TIRL framework, i.e. the developer who writes the stage-1 registration script and decides what transformations should be added to the chain for an automated registration.) The user who runs the stage-1 script only has to provide the resolution with the configurations; the weighted centre is calculated automatically when the transformation is added to the chain. In the internal ("TIRL-managed") part, transformations are both added and configured automatically by TIRL, as we will see in the case of the c_H. The c_B chain has no *internal* part.

3.6.3 Deconstructing the transformation chain of the histology image

The chain of the histology image (c_H) is an injection that maps the pixel coordinates of the histology image into physical space:

$$c_H : D_0 \rightarrow S_4 : \mathbf{x}'_H = c_H(\mathbf{x}_H) \tag{3.3}$$

It is more complex than c_B, in keeping with the fact that the histology image is the more distorted one of the inputs, and we aim to undo these by a series of elementary transformations. Figure 3.4 shows the chain subdivided into one internal and three external subunits.

1. The first subunit is the *internal* part of c_H. It is termed internal, because it is created automatically by TIRL when H is downsampled and padded during the pre-processing stage. In general, internal transformations cancel the side effects that image operations have on the image coordinates. In this specific case, the internal part comprises two transformations. The first one is a translation (*Trim*) that zeroes the offset of the top-left pixel of the downsampled histology image that was introduced by the padding. If the padding was 50 pixels along the vertical dimension, and 60 pixel along the horizontal dimension, the parameters of this transformation would be $(-50, -60)$. The second is a scaling transformation (*Upscale*), that maps the offset-compensated pixel coordinates of the downsampled image to the pixel coordinates of the original histology image. For example, if the downsampling factor was 8, the $(0, 1)$ offset-compensated pixel of the downsampled image would map to $(0, 8)$, i.e. the 8th pixel in the the top row of the original histology image. The net effect of the internal part of the transformation chain is that the points of a downsampled and padded histology image are mapped to the pixel coordinates of the histology image, in the form that it was loaded from the SVS file. This space, the pixel space of the original histology image is denoted by S_1.

   The advantage of the auto-generated internal subunit is that the later parts of the chain only depend on S_1, and not the padding width or the sampling level. Consequently, it is very easy to split c_H and reuse the later parts only to transform point sets that are defined on the original histology image, or other

images that have been previously registered to the original histology image. This illustrates the general flexibility of TIRL, that is to make transformations between image spaces as intuitive as possible: any subchain that maps to a certain domain can be combined with any other subchain that maps from the same domain. Table 3.1 gives a description of each intermediate domain that is indicated in Figure 3.4.

Space	Description
D_0	H in a padded, coarse (50 µm/px) pixel space, origin: top-left corner
S_1	H in its original pixel space (8 µm/px), origin: top-left corner
S_2	H in metric space, origin: weighted image centre
D_3	H in metric space after linear registration
S_4	H in metric space after all of registration
	B in metric space, origin: weighted image centre
D_5	B in its original voxel space (50 µm/px) origin: top-left corner

Table 3.1: Space and image domain definitions related to the stage-1 transformation chain.

2. The second subunit (*init*) is in the *external* part of c_H, meaning that it is user-defined. It is the direct equivalent of the c_B chain, and it initialises the histology image in metric space. It comprises a scaling transformation (*Resolution*), and a translation component (*Centre*). Both S_2 and S_4 are embedded in metric space and their origins coincide. Consequently, the initialisation addresses the issues of different resolutions and disjoint ROIs.

3. The third subunit (*linear*) defines the chain of elementary linear transformations that are optimised during stage 1: a rotation (*Rotation*) (about the origin of S_2, i.e. the weighted image centre), followed by isotropic scaling (*Iso-scale*), a translation (*Translation*), and finally a full 6-DOF affine transformation (*Affine*). The parameter redundancy with the previous transformations is mitigated by optimising the affine transformation separately, while keeping the other transformation parameters constant. The order of the elementary transformations in the linear subchain is a deliberate choice. The rotation and isotropic scaling come before translations to ensure that rotations are

defined about the weighted centre of the image, and scale changes do not affect the offset of the image from the origin. As long as the scaling is isotropic, it is interchangeable with the rotation, as it will not introduce shears. The affine transformation is applied at the end to refine any of the rigid and scale parameters, and add shears only if those were insufficient. The linear subunit maps the initialised histology image coordinates further in metric space, as close to the points of the tissue block photograph as reasonably possible with global distortions. The coordinate system at the end of the linear subunit coincides with the domain where the non-linear deformations are defined: D_3.

4. The fourth subunit (*non-linear*) consists of a single non-linear transformation. It is parametrised as a dense displacement field (`TxDisplacementField`), that assigns a displacement vector to each pixel of the downsampled histology image (H). However, as the image coordinates are now in metric space, both the locations of the displacement vectors and the displacement vectors themselves must be defined in the same space where the image points are. Therefore, D_3 is instantiated from the `Domain` object of D_0 (so that they share the same pixel grid), but with all the antecedent transformations of c_H to map to coordinates after the linear transformations. To reduce the computational cost, the displacement vectors are estimated directly in D_3 space, without having to map the vector definitions back to pixel space on every iteration.

Altogether, the optimised part of the c_H chain performs an affine mapping of the physical coordinates of the initialised original histology image (S_2), followed by a non-linear refinement. Using homogeneous coordinates, the non-linear mapping from the best affine alignment (D_3):

$$\mathbf{x}'_H = \mathbf{x}_H^{(l)} + \mathbf{u}(x_H^{(l)}) \tag{3.4}$$

where $\mathbf{x}'_H \in S_4$ and $\mathbf{u}$ denotes the displacement vector at a given point of the deformation field.

3.6.4 Pixel-to-pixel mapping by the unified stage-1 transformation chain

In the previous sections we have seen how the initialisation into metric spaces S_2 and S_4 addressed the issues of different image resolutions and disjoint ROIs. To calculate the similarity of the transformed histology image and the tissue block photo, these two must be represented by the same number of pixels. To achieve this, we combine the previously introduced transformation chains c_H and c_B to create a unified stage-1 transformation chain, and establish pixel-to-pixel correspondence between the histology image and the tissue block photograph by interpolation.

We start by defining the unified stage-1 transformation chain (c_1) as:

$$c_1 = (c_B)^{-1} \circ c_H \tag{3.5}$$

where c_B^{-1} is the inverse of c_B. The resultant c_1 is a forward mapping from the pixel domain of the histology image to the pixel domain of the tissue block photograph: $c_1 : D_0 \rightarrow D_5$. Given the transformation parameters of the unified chain ($\mathbf{p}$), TIRL maps D_0 onto the pixel space of the tissue block photo, and *evaluates* B at the respective locations to create an alternative tissue block image (B_1) that is in one-to-one correspondence with the pixels of the histology image:

$$B_1(\mathbf{x}_H, \mathbf{p}) = B \circ c_1(\mathbf{x}_H, \mathbf{p}) \tag{3.6}$$

The values of B_1 are obtained by bilinear interpolation of the image values in B at the transformed locations.

3.6.5 Optimisation substages

Once the stage-1 transformation chain is set up and properly initialised, the optimisation of the linear and non-linear subchains begins. The goal of the optimisation is to find the parameter vector $\mathbf{p}$ that provides the best match between the transformed histology image and the tissue block photograph. For a large number of slides (>120 blocks) it would be impractical to rely on the registration of manually annotated anatomical landmarks, therefore we use the MIND metric (see

Section 2.5) to assess image similarity and define the following objective function to be minimised (in discrete form, in keeping with the actual implementation):

$$C_1(\mathbf{p}) = \frac{1}{2} \sum_{\mathbf{x}_H \in D_0} \|\mathcal{M} \circ H(\mathbf{x}_H) - \mathcal{M} \circ B_1(\mathbf{x}_H, \mathbf{p})\|_2^2 \tag{3.7}$$

where $\mathbf{x}_H$ denote histology pixel coordinates, and $\mathcal{M}$ is the MIND operation.

Not all parameters of the unified stage-1 chain are optimised at once. By optimising transformations in the order of increasing degrees of freedom, previously optimised coarser transformations provide a more suitable initialisation for finer transformations that are optimised later. The registration therefore proceeds from estimating the most influential parameter (rotation) towards optimising the least influential ones (e.g. one local displacement vector component). The final transformation parameters are obtained through three linear and one non-linear optimisation substages, which are discussed below.

1. **Rotation search.** The initialisation does not change the relative orientation of the input images, although we found this highly variable in our post-mortem dataset. A relative orientation of 90° or 180° is not uncommon, as the tissue section might need a 90° rotation to fully fit on the narrow microscope slide (25 x 76 mm), after which it might get mounted in an upright or upside down orientation. Local optimisation methods are not well suited to estimate rotations that are this large, because the optimisation would need to cross over rotation states that are associated with a high cost (Figure 3.5). This necessitates a rotation search.

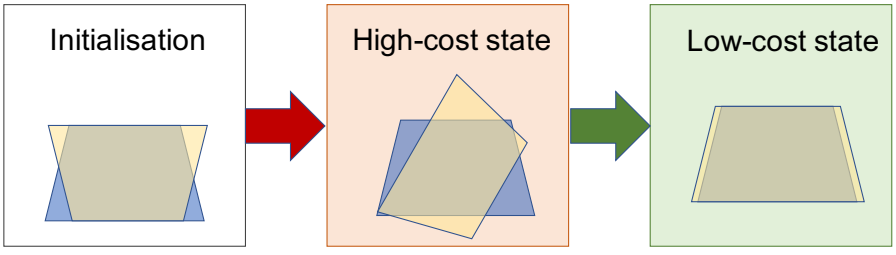

Figure 3.5: Large orientation differences cannot be reliably compensated by local optimisation, as the state associated with the lowest cost is only accessible through states that are associated with higher-than-initial cost.

The input images are downsampled to 0.5 mm resolution to speed up computations and reduce the sensitivity to misalignments. Then the histology image is initialised by 30° increments in the $[-180°, 180)$ range, and the cost is calculated for each angle. We select the 3 best rotation angles, and run the next optimisation step (similarity transformation) from each initialisation. The optimised rotation angle is then taken from the best result.

2. **Similarity transformation.** The parameters optimised in this substage are the rotation angle, the isotropic scaling factor, and the translation parameters. The scale factor is expected to be close to 1, only compensating for the inaccuracy of the tissue block photograph resolution, which comes from manual input. This substage employs the gradient-free BOBYQA optimiser [55], which allows for the specification of hard bounds on each parameter. Note that this substage is executed four times. Thrice as part of the rotation search, and one final time, after the best initial rotation angle was found. During the first three executions, the rotation parameter is constrained within a $\pm 15°$ range around the initial coarse value, and the optimisation is carried out at 0.5 mm resolution. When the optimisation runs for the final time, the rotation range is $\pm 15°$ around the previously optimised rotation angle, and the optimisation is carried out on a resolution pyramid [91] (0.8 mm, 0.4 mm, and 0.2 mm). Smoothing may optionally be applied to the images at each resolution level in TIRL, but we managed to obtain satisfactory results without this option.

3. **Affine transformation.** This substage optimises the 6 parameters of the *Affine* transformation of the chain. The affine matrix introduces shear, which is particularly relevant to histology images, as the thin section of tissue may also deform this way when it is mounted on the microscope slide. Optimising shear with the rest of the affine parameters leaves space for small adjustments on previously estimated rigid parameters (and the scale factor), which are held constant during this substage to prevent parameter redundancy. The

optimisation scheme is BOBYQA, where the bounds for the diagonal elements of the affine matrix are ± 0.05, ± 0.2 for the off-diagonal elements, and ± 1 mm for the translation components.

4. **Non-linear transformation.** The non-linear registration used by stage 1 is based on the concept of *diffusion registration* by Modersitzki et al [92]. Diffusion registration ensures that the calculated deformation field is smoothly varying by adding a regularisation term to the cost function that penalises large spatial gradients in the deformation field:

$$\begin{aligned} C_2(\mathbf{u}, \alpha) = \frac{1}{2} \sum_{\mathbf{x} \in D_0} \| \mathcal{M} \circ H(\mathbf{x}) - \mathcal{M} \circ B(\mathbf{M}\mathbf{x} + \mathbf{u}(\mathbf{x})) \|_2^2 + \\ \alpha \cdot \frac{1}{2} \sum_{\mathbf{x} \in D_0} \sum_{i=1}^{2} \| \nabla u_i(\mathbf{x}) \|_2^2 \end{aligned} \tag{3.8}$$

where $\mathbf{x} = \mathbf{x}_H$ (pixel coordinates of the histology image), $\mathcal{M}$ is the MIND operation, and $\alpha > 0$ is a weighting factor for the regularisation term that needs to be determined empirically. Note, that the first term in 3.8 is identical to C_1 in 3.7. As the non-linear transformation parameters are the only parameters optimised in this substage, these are represented in 3.8 as a vector field, such that $u_i(\mathbf{x})$ refers to the i-th component of the displacement vector, $\mathbf{u}(\mathbf{x})$, at the location $\mathbf{x}$. All previously optimised linear transformations are collectively represented in 3.8 by an affine mapping ($\mathbf{M}$).

This cost function can be efficiently optimised by techniques that Heinrich et al [50] pointed out. The first of these is to rewrite the discrete expression in 3.8 using a bilinear form $\sum_{\mathbf{x}} \mathbf{f}(\mathbf{u}(\mathbf{x}))^\top \mathbf{f}(\mathbf{u}(\mathbf{x}))$. For this, we define the following two vectors for the dissimilarity and regularisation terms, respectively:

$$\mathbf{f}_D(\mathbf{u}(\mathbf{x})) = \mathcal{M} \circ H(\mathbf{x}) - \mathcal{M} \circ B(\mathbf{M}\mathbf{x} + \mathbf{u}(\mathbf{x})) \tag{3.9}$$

$$\mathbf{f}_R(\mathbf{u}(\mathbf{x})) = \begin{bmatrix} \frac{\partial u_1}{\partial x} & \frac{\partial u_1}{\partial y} & \frac{\partial u_2}{\partial x} & \frac{\partial u_2}{\partial y} \end{bmatrix}^\top \tag{3.10}$$

Substituting 3.9 and 3.10 into 3.8, we get the following expression for the cost function:

$$\begin{aligned} C_2(\mathbf{u},\alpha) &= \frac{1}{2}\sum_{\mathbf{x}\in D_0}\mathbf{f}_D(\mathbf{x})^\top\mathbf{f}_D(\mathbf{x}) + \frac{\alpha}{2}\sum_{\mathbf{x}\in D_0}\mathbf{f}_R(\mathbf{x})^\top\mathbf{f}_R(\mathbf{x}) = \\ &= \frac{1}{2}\sum_{\mathbf{x}\in D_0}\begin{bmatrix}\mathbf{f}_D(\mathbf{x})\\ \sqrt{\alpha}\mathbf{f}_R(\mathbf{x})\end{bmatrix}^\top\begin{bmatrix}\mathbf{f}_D(\mathbf{x})\\ \sqrt{\alpha}\mathbf{f}_R(\mathbf{x})\end{bmatrix} = \frac{1}{2}\sum_{\mathbf{x}\in D_0}\mathbf{f}(\mathbf{u}(\mathbf{x}))^\top\mathbf{f}(\mathbf{u}(\mathbf{x})) \end{aligned} \tag{3.11}$$

The power of the idea that the entire cost function in 3.8 can be expressed in bilinear form is that it can be minimised with respect to $\mathbf{u}$ using Gauss–Newton optimisation, where an initial estimate of the deformation field $\mathbf{u}^{(0)}(\mathbf{x})$ is iteratively modified by $\mathbf{h}^{(k)}(\mathbf{x})$, until the updates vanish or a certain number of iterations is reached:

$$\mathbf{u}^{(k+1)}(\mathbf{x}) = \mathbf{u}^{(k)}(\mathbf{x}) + \mathbf{h}^{(k)}(\mathbf{x}) \tag{3.12}$$

In the implementation of this substage, the iterations are set to terminate after any of the following criteria is met:

(a) 20 iterations completed

(b) The L2 norm of the parameter updates (that is, $\|\mathbf{u}^{(k+1)}(\mathbf{x}) - \mathbf{u}^{(k)}(\mathbf{x})\|$) are less than 0.1 pixel.

(c) The L2 norm of the parameter updates are less than 1% of the L2 norm of the current transformation parameters, $\mathbf{u}^{(k)}(\mathbf{x})$.

In the Gauss–Newton method, updates are calculated by solving the following linear system [52] for all pixels:

$$\left(\mathbf{J}^\top\mathbf{J}\right)\mathbf{h}^{(k+1)}(\mathbf{x}) = -\mathbf{J}^\top\mathbf{f}(\mathbf{u}^{(k)}(\mathbf{x})) \tag{3.13}$$

where $\mathbf{J}$ is the derivative of $\mathbf{f}$ with respect to $\mathbf{u}$. Heinrich et al [50] apply the Gauss–Newton formula in 3.13 to the regularised cost expression in 3.11 and derive that $\mathbf{h}^{(k+1)}(\mathbf{x})$ can be calculated in each iteration by simultaneously solving the following linear system for all pixels, starting from the initial estimate $\mathbf{h}^{(0)}(\mathbf{x}) = 0$ for all $\mathbf{x}$:

$$\left(\nabla\mathbf{f}_D^\top\nabla\mathbf{f}_D - \alpha\Delta\right)\mathbf{h}^{(k+1)}(\mathbf{x}) = -\left(\nabla\mathbf{f}_D^\top\mathbf{f}_D - \alpha\Delta\mathbf{u}^{(k)}(\mathbf{x})\right) \tag{3.14}$$

where $\nabla \mathbf{f}_D = \begin{bmatrix} \frac{\delta \mathbf{f}_D}{\delta u_1} & \frac{\delta \mathbf{f}_D}{\delta u_2} \end{bmatrix}^\top$ and Δ denotes the discrete Laplace operator, i.e. $\Delta \mathbf{u}(\mathbf{x}) = \nabla \cdot \nabla \mathbf{u}(\mathbf{x})$. Note that it is difficult to compute the exact value of $\nabla \mathbf{f}_D$, since $\mathbf{f}_D$ is not linear with respect to $\mathbf{u}$. Instead, we can use the first-order Taylor-expansion of $\mathbf{f}_D$ to obtain an approximate value of $\nabla \mathbf{f}_D$ at each iteration. First we evaluate the tissue block photograph on the domain of the histology image using the current estimate of the displacement field, and calculate the MIND representation of the result:

$$B_2(\mathbf{x}) = \mathcal{M} \circ B(\mathbf{M}\mathbf{x} + \mathbf{u}(\mathbf{x})) \tag{3.15}$$

Then do Taylor-expansion on $\mathbf{f}_D$:

$$\mathbf{f}_D(\mathbf{u}(\mathbf{x}) + \mathbf{h}(\mathbf{x})) \approx \mathcal{M} \circ H(\mathbf{x}) - B_2(\mathbf{x}) - \nabla B_2(\mathbf{x})^\top \mathbf{h}(\mathbf{x}) \tag{3.16}$$

From which it follows that:

$$\nabla \mathbf{f}_D \approx -\nabla B_2 \tag{3.17}$$

Since B_2 is an image, it can be seen that the $\nabla \mathbf{f}_D^\top \nabla \mathbf{f}_D$ expression is a 2-D structure tensor.

TIRL solves 3.14 simultaneously for all $N = |D_0|$ pixels when the respective optimiser object (`GNOptimiserDiffusion`) is called. For this purpose, a linear system of equations, encompassing all pixels, is defined as a matrix equation. Current estimates of the displacement field ($\mathbf{u}^{(k)}$) are flattened such that displacements along x (an N-dimensional vector) are appended by displacements along y to create a $2N$-dimensional vector ($\mathbf{U}$). From the four components of the structure tensor ($(\nabla \mathbf{f}_D(\mathbf{x}))^\top \nabla \mathbf{f}_D(\mathbf{x})$), four separate $N \times N$ diagonal matrices ($\mathbf{XX}, \mathbf{XY}, \mathbf{YY}$) are created and assembled into a $2N \times 2N$ block matrix ($\mathbf{A}_{st}$):

$$\mathbf{A}_{st} = \begin{bmatrix} \mathbf{XX} & \mathbf{XY} \\ \mathbf{XY} & \mathbf{YY} \end{bmatrix} \tag{3.18}$$

An $N \times N$ matrix ($\mathbf{L}$) is created from the discretisation of the Laplace operator on the $m \times n$ large pixel domain of the histology image (D_0) using a 5-point

stencil (see 3.19). Note that $\mathbf{L}$ is sparse, having 4's on its main diagonal, -1's on its $k = 1, n$ supradiagonals and -1's on the $k = -1, -m$ subdiagonals. All other elements are zero.

$$\mathbf{L} = \begin{bmatrix} 4 & -1 & 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ -1 & 4 & -1 & 0 & \dots & 0 & -1 & 0 & \vdots \\ 0 & -1 & 4 & \ddots & 0 & \dots & 0 & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & -1 & \ddots & & \ddots & -1 \\ 0 & & 0 & -1 & 4 & -1 & 0 & & 0 \\ -1 & \ddots & & \ddots & -1 & \ddots & \ddots & 0 & \vdots \\ 0 & \ddots & 0 & \dots & 0 & \ddots & 4 & -1 & 0 \\ \vdots & 0 & -1 & 0 & \dots & 0 & -1 & 4 & -1 \\ 0 & \dots & 0 & -1 & 0 & \dots & 0 & -1 & 4 \end{bmatrix} \tag{3.19}$$

The expression $\nabla \mathbf{f}_D^\top \mathbf{f}_D$ is evaluated directly for each pixel and the results are stored in a $2N$-dimensional vector $\mathbf{F}$.

Based on the above definitions and 3.14, it is the linear system below that TIRL solves at each iteration until one of the stopping criteria is met, using the (`minres`) iterative solver of the SciPy package:

$$\left(\mathbf{A}_{st} - \begin{bmatrix} \alpha\mathbf{L} & \mathbf{0} \\ \mathbf{0} & \alpha\mathbf{L} \end{bmatrix}\right) \mathbf{H}^{(k+1)} = -\left(\mathbf{F} - \begin{bmatrix} \alpha\mathbf{L} & \mathbf{0} \\ \mathbf{0} & \alpha\mathbf{L} \end{bmatrix} \mathbf{U}^{(k)}\right) \tag{3.20}$$

where $\mathbf{H}^{(k+1)}$ is the $2N$-dimensional vector that contains the updates to the displacement vectors for every pixel.

TIRL also allows the diffusion registration to be repeated at multiple resolution levels and with additional smoothing of the images. By default, stage 1 carries out diffusion registration at 0.8 mm, 0.4 mm, and 0.2 mm resolutions, which we found to be a good compromise between registration accuracy and speed. No additional smoothing was applied to the images at any of those resolution levels. The field is upsampled between levels using bilinear interpolation.

The multiresolution optimisation strategy accounts for large bulk deformations that cannot be compensated by the transformations of the whole image (e.g.

affine). These are areas in the histology image, for example where a piece of the tissue moved as a whole.

3.7 Evaluating the accuracy of the registration

To test the accuracy of the stage-1 pipeline, we carried out registrations on 14 corresponding histology and tissue block images of the corpus callosum (CC), and further 14 pairs of images from the hippocampus (HC). The two regions were chosen based on the availability of histology images from the ongoing post-mortem study, as well as to be representative of the two variants of the human cerebral cortex: the isocortex (in CC) and the allocortex (in HC). After observing the images, we identified the gray-white matter boundary as the most ubiquitous and most characteristic feature in our images, hence we based our measurements of accuracy on this boundary.

All 28 pairs of images were pre-processed (Section 3.5) and registered (Section 3.6) using the same pipeline and configurations. The gray-white matter boundary was traced by hand in all original histology images and tissue block photographs using the ImageJ software. While drawing contours, the histology and the tissue block photos were open side-by-side to aid the recognition of the contour in areas where it became too dim in either of the images. While this approach might impose a small bias on the segmentations, it is reasonable to assume that contours have similar shapes in both images, and the anticipated gain in the registration accuracy from this bias is much smaller than the anticipated loss of accuracy in the case of a misidentified contour. Care was taken to match the locations of the start and end points of the contour between the histology and the tissue block image to achieve a rough pairwise correspondence between the intermediate points of the contours. Each contour was defined by approximately 200 points, which were upsampled to a constant 2000 points using cubic spline interpolation.

To make contours comparable, the tissue block contour was transformed to the pixel space of the original histology image (S_2) by using c_B and the inverse of the c_H-subchain (cf. Figure 3.4). The accuracy of the registration was quantified by

the *mean contour distance* (MCD) in millimetres. The MCD was measured by recording the distance between each point on one contour and the closest point on the other contour, repeating the process for both contours, and taking the mean of all 4000 measurements. As this approach looks for the closest point, no strict pairwise correspondence is expected between points with the same index. On the other hand, only taking the closest point could lead to the underestimation of the registration error if two distant parts of the curves overlapped. The search range for the closest point was therefore confined to ± 250 points (upstream and downstream) relative to the point with the corresponding index, such that accidental overlaps between non-equivalent parts of the curves were attributed correctly.

As the MCD is affected by the regularity and the accuracy of the annotations, care was taken to minimise these errors. Compared with the Dice coefficient, which also depends on the accuracy of segmentations, an important advantage of the MCD is that it gives an estimate of the registration error in millimetres, which is more intuitive to understand.

Figure 3.6 shows the evolution of the MCD on a PLP-stained section of the corpus callosum through the substages of stage-1 registration. As evidenced by the contour distances, each substage improves the alignment. Despite the obvious difference in the scale and orientation of the input images (panels A, B), the rotation search (C) can find the correct orientation of the images. The similarity transform (D) refines the scale and translation parameters, as well as the affine stage (E) that starts from the best rigid alignment. The largest improvement is made by the non-linear stage (F). The final registration error (0.15 mm) is in close correspondence with the resolution at which the displacement field was optimised (0.2 mm). This optimisation was carried out with $\alpha = 0.3$ regularisation weight.

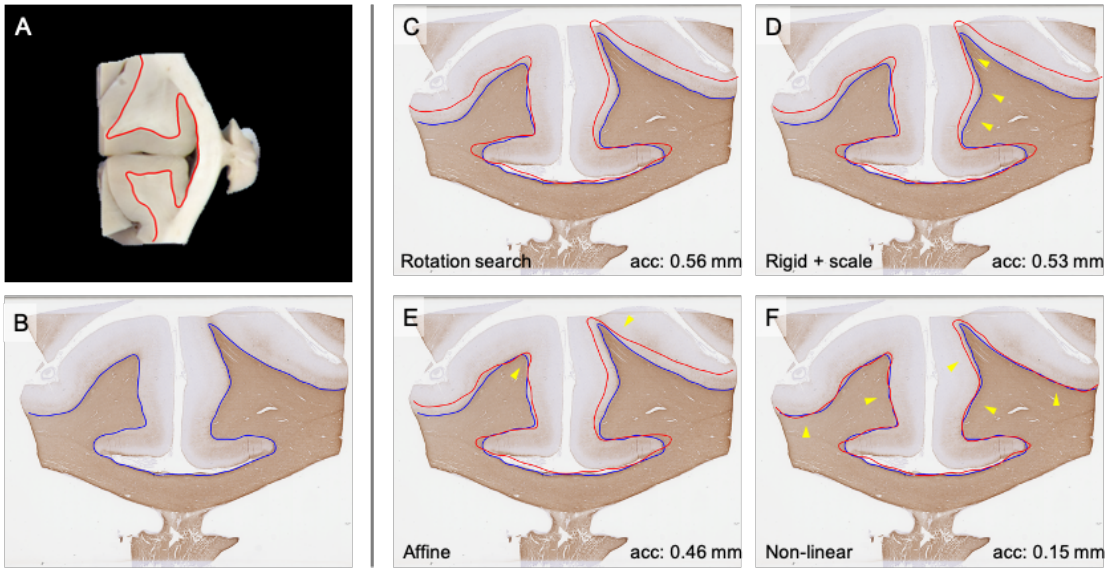

Figure 3.6: Stage-1 accuracy is evaluated by mean contour distance. Contours (*red*, *blue*) are hand-drawn at the gray-white matter interface on the tissue block photograph **(A)** and the histology image **(B)**. **(C–F)** show a gradually improving overlap between the contours after each substage of the registration (see main text). The *yellow arrowheads* indicate the largest changes relative to the previous substage. As indicated by the mean contour distances, the non-linear stage is the largest contributor to registration accuracy. Note: only the contours are mapped onto the histology image – the image itself is identical in all panels of the figure.

3.8 Exploring the effects of regularisation

The non-linear registration substage of stage 1 defines a metaparameter, the regularisation weight (α) (cf. Equation 3.8), which must be set in advance of the registration. The larger α is, the smoother the deformation field will become, at the expense of more residual misalignment of the images. The optimum regularisation weight must be determined empirically for the particular registration task.

Registrations on the 14 hippocampal and 14 callosal sections were run with varying degrees of regularisation ($\alpha \in \{0.2, 0.3, 0.4, 0.6, 0.8, 1.0\}$) to determine the optimal α for each series. Figure 3.7 shows a comparison of the estimated non-linear deformations on a hippocampal section for the different values of α. The range of Jacobian determinants and the MCD were calculated for each case. With $\alpha = 0.2$ the Jacobian range is -9.5-13.4, indicating excessive shinkage and expansion of the pixels, which is also evident by observing the large, fluid-like distortions of the

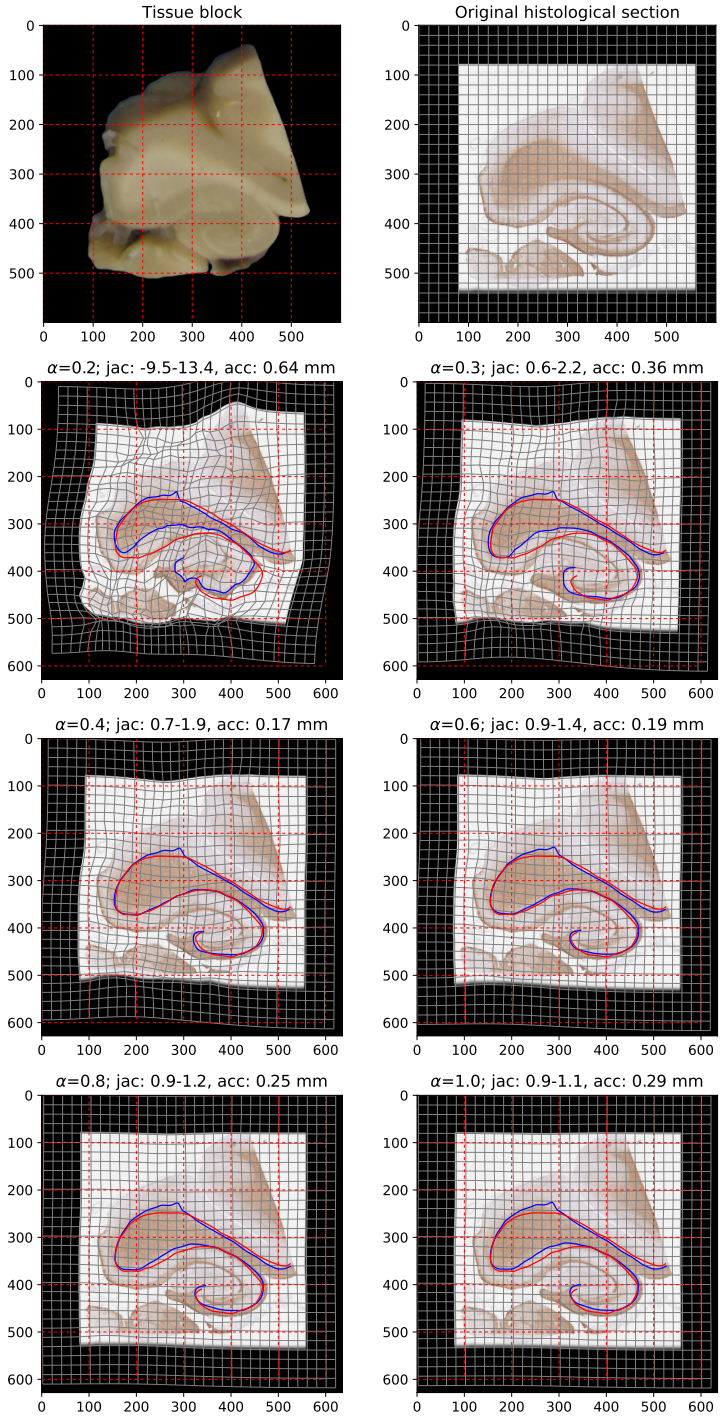

Figure 3.7: Non-linear registration of a hippocampal section after affine initialisation using varying degrees of regularisation ($0.2 \leq \alpha \leq 1.0$). All images are show on the domain of the original tissue block photograph (D_5). The histology image in the top row is shown after the affine substage was completed. *Red grid lines* are provided for comparison, *gray grid lines* show the estimated deformations of the histology image. Jacobian ranges (*jac*) and mean contour distances (*acc*) are reported for each case. The *blue contour* was obtained by transforming the contour from the original histology image, the *red contour* was drawn on the original tissue block photograph to highlight the accuracy of the alignment.

histology grid (gray grid). Negative determinants imply that some parts of the image were lost due to folding, and the corresponding deformation field is non-invertible. For $\alpha > 0.2$ the Jacobian determinants are positive, gradually closing around unity as the value of α increases. At $\alpha = 0.3$ the deformations already appear physically realistic, but there is not enough correlation between neighbouring displacement vectors to account for the bulk deformation of the hippocampus as a whole, leading to a large misalignment of the subiculum and the cornu ammonis. The contour distances reveal that in this particular case, the most accurate result (MCD = 0.17 mm) was obtained using $\alpha = 0.4$. For $\alpha = 0.6$, the registration is still accurate at the resolution that the deformation field was optimised at (0.2 mm). For larger values of α, the gray grid lines appear barely deformed relative to the affine substage. While the bulk deformation of the hippocampus is effectively captured (contours align around the cornu ammonis), smaller deformations of the parahippocampal gyrus are not compensated effectively.

To test how well the optimum value of α generalises from one image to several images, including the CC and the HC, we measured the MCD on all 28 pairs of images for different values of α. Additionally, the MCD was also measured for the linear substages of the pipeline. For a fair comparison between different regularisation weights, the non-linear substage was run from the same affine initialisation (D_3) for all values of α.

Figures 3.8 and 3.9 show the distribution of MCD for each substage and regularisation weight on the CC and the HC sections, respectively. In both series the linear substages constituted a gradual improvement in the alignment of the images, and the non-linear substage significantly improved the accuracy of the registration, confirming that the tissue sections were distorted by histological processing.

For the CC, the difference between the MCD after the rotation search and the similarity transform was minimal, and the affine substage seemed to have a stronger influence on the images that were highly misaligned after the previous stages. An inspection of these images revealed that shears were interfering with the correct estimation of the rotation and the scale factor in the first substages,

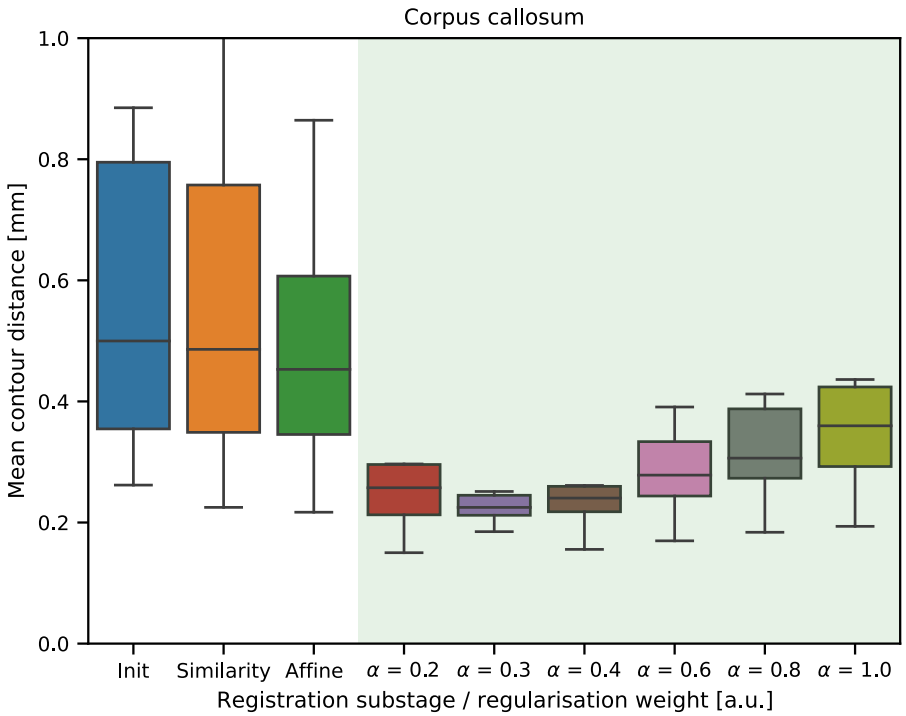

Figure 3.8: Accuracy of stage-1 registration measured by mean contour distance on 14 callosal sections with manually-defined ground truth. Accuracy is reported separately for each linear substage (*white background*), as well as the non-linear substage, using different values of the regularisation parameter α (*green background*). Horizontal lines represent the median, boxes show the interquartile range, and the bars span the whole range of the data.

which were successfully compensated by the affine substage. As expected, the α coefficient had an optimum in the test range (0.2-1.0). The optimum was found at $\alpha = 0.3$ with the lowest median MCD (0.22 mm), but similar registration errors were observed with $\alpha = 0.4$ (0.23 mm). At the highest regularisation value of $\alpha = 1.0$ the alignment was still noticeably better compared with the result of the affine substage, especially for sections with larger initial misaligment (as evidenced by shrinkage of the interquartile ranges and the top bar). This is in keeping with the existence of bulk deformations (e.g. a slight bending of a gyrus) in the mounted tissue slice that cannot be compensated by global transformations, but are still captured by the highly regularised non-linear substage. However, high regularisation cannot account for the more locally confined distortions of the tissue (e.g. anisotropic stretching). The Jacobian determinants were positive throughout the image for all cases where $\alpha > 0.2$.

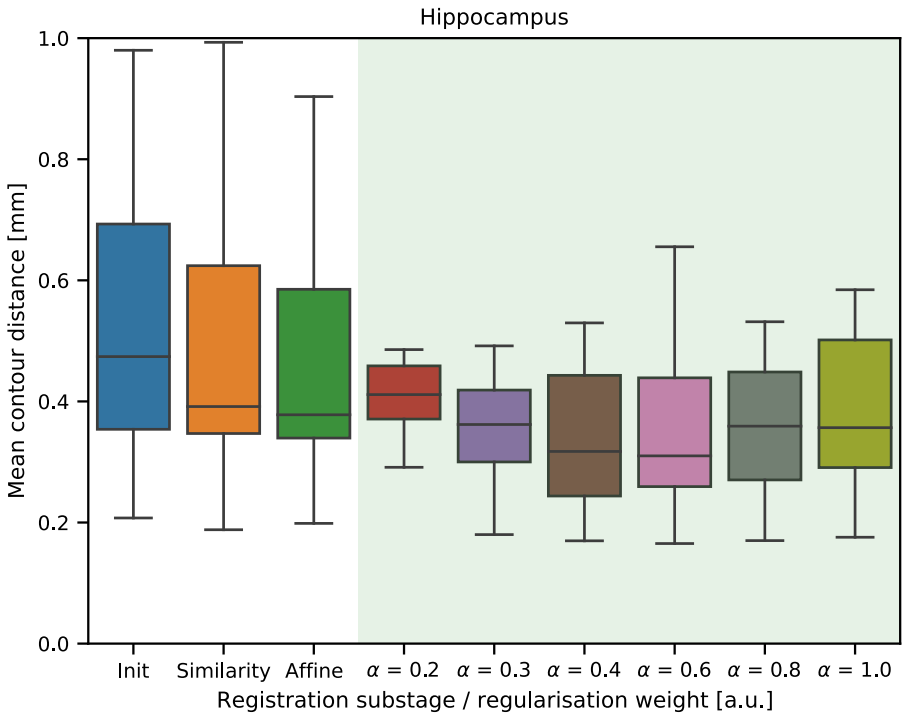

Figure 3.9: Accuracy of stage-1 registration measured by mean contour distance on 14 hippocampal sections with manually-defined ground truth. Accuracy is reported separately for each linear substage (*white background*), as well as the non-linear substage, using different values of the regularisation parameter α (*green backgorund*). Horizontal lines represent the median, boxes show the interquartile range, and the bars span the whole range of the data.

For the HC, the similarity transform was associated with a larger relative improvement of the registration accuracy than in the case of the CC. A likely explanation is the lower contrast in the tissue block photographs plus the high rotational symmetry, increasing the chances of incorrect rotation estimation and subsequent scaling errors. Given the smaller size of the blocks, shears contribute less to the initial registration error, which explains the moderate improvement by the affine substage. For the non-linear substage, the optimal regularisation weight was found to be $\alpha = 0.6$. Compared to the CC, the registration errors for the HC were generally larger in all substages. This was attributed to a combination of multiple factors:

1. The general low contrast in the tissue block photographs of the HC impacted the definition of high-quality contours.

2. The low contrast also impacted correspondence of image features between the histology and the block photo that the MIND-based registration depends on.

3. The tissue blocks of the HC presented with visible side surfaces in the photographs, which were unmatched in the histology image, and these were not masked from the algorithm.

4. The HC sections were more severely affected by changes of the anatomy as tissue sections were taken from a greater depth within the blocks. (This phenomenon is discussed in greater detail in Chapter 6)

Despite the seemingly less accurate alignment of the HC sections, it should be noted that the results were visually satisfactory on observation, and that the measured registration errors are smaller than the typical voxel sizes of high-resolution MRI scans ($\leq$0.5 mm), which should be further improved by addressing some or all of the factors listed above.

3.9 Registering sections with different stains

To test how well the stage-1 pipeline generalises for different histology stains, we tested the registration of a section from the orbitofrontal cortex, that was stained by 6 different methods (two tincture stains: H&E, LFB+PAS, and four immunostains: Ferritin, PLP, CD68, pTDP43). Half of the stains (Ferritin, PLP, LFB+PAS) exhibited good gray-white matter contrast, although the contrast was barely noticeable at the 8 µm resolution. The images were pre-processed and registered with the same stage-1 protocol and configurations as outlined in Sections 3.5 and 3.6. The regularisation weight was set to 0.3, in line with its optimal value on the CC sections. A single contour was drawn on the gray-white matter boundary of the tissue block photograph, and transformed into the pixel space of the original histology images (S_2). The Jacobian determinants were calculated and plotted for the deformation field corresponding to each stain. The results are shown in Figures 3.10 and 3.11.

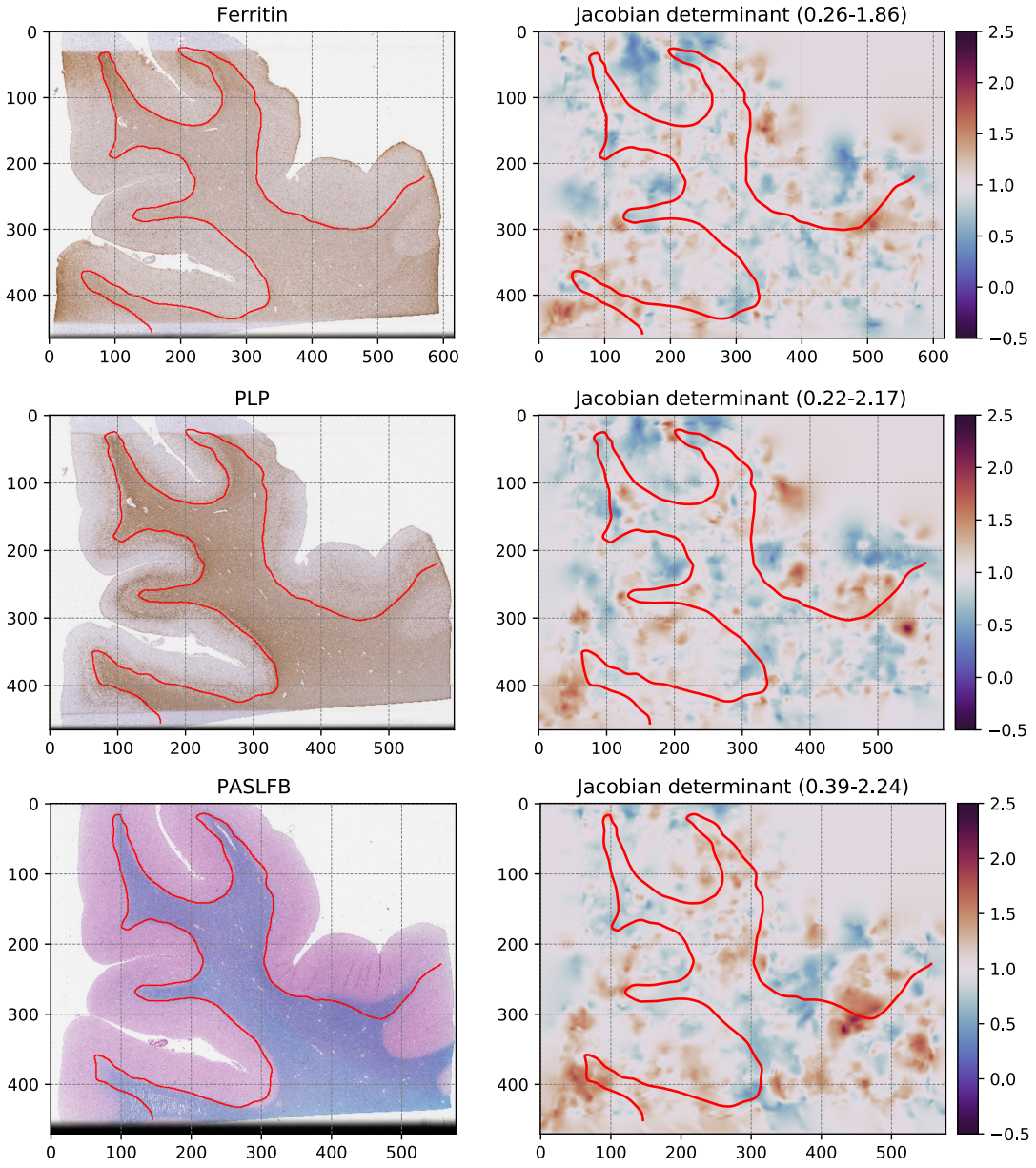

Figure 3.10: Stage-1 results on a section of the orbitofrontal cortex immuno-stained for PLP, Ferritin and tincture-stained with LFB+PAS. All these stains exhibit good gray-white matter contrast. The *red contour* corresponds to the transformed gray-white matter boundary of the tissue block photograph. Jacobian determinants are in a plausible range, and the Jacobian maps are qualitatively similar.

Both figures display a good general alignment between the features of the histology images and the transformed gray-white matter contour of the corresponding tissue block photograph, irrespective of the gray-white matter contrast of the images. The range of Jacobian determinants within the area of the tissue sections are

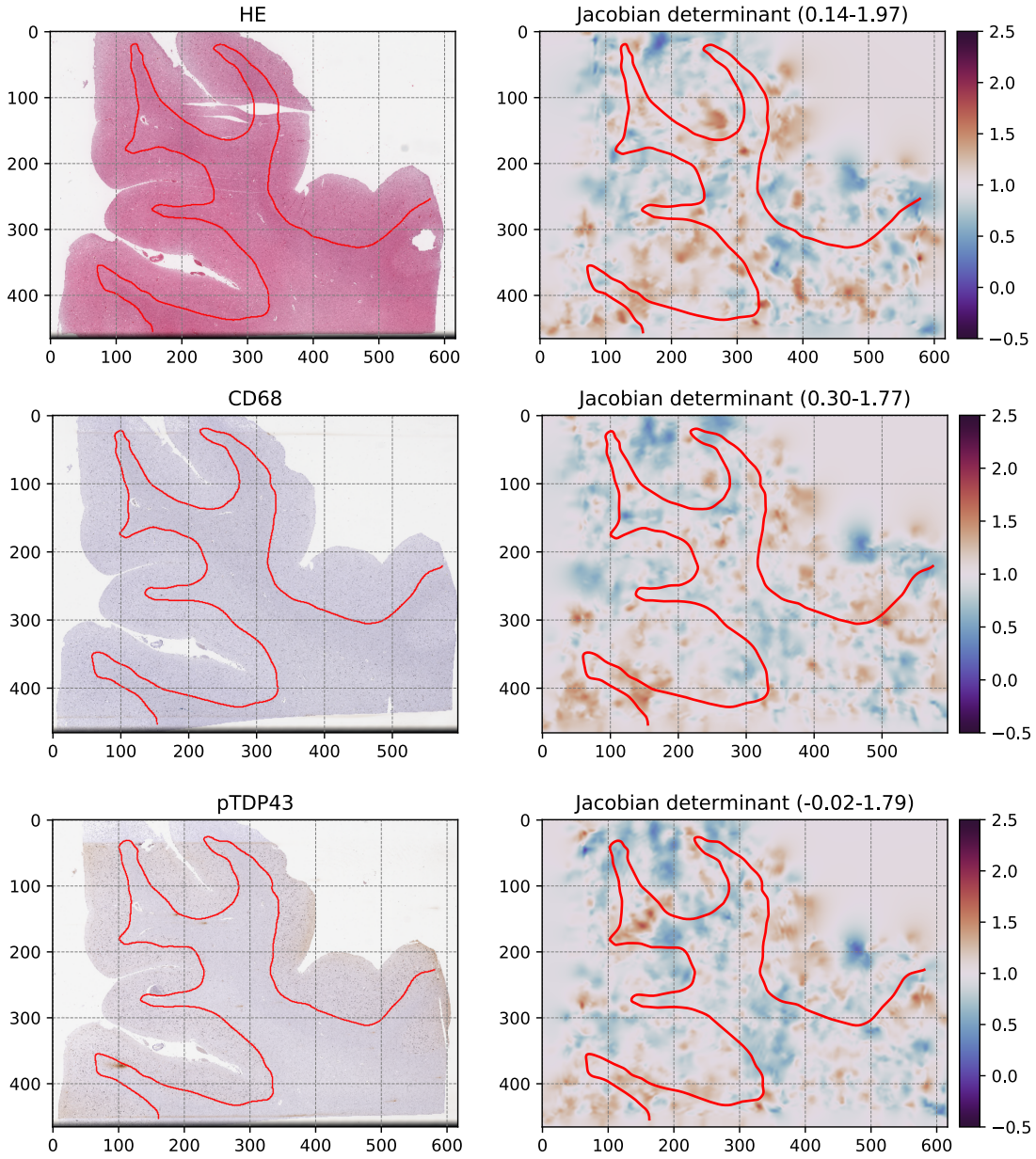

Figure 3.11: Stage-1 results on a section of the orbitofrontal cortex tincture-stained with H&E, and immuno-stained for CD68 and pTDP-43. These stains exhibit a dim contrast between the gray and the white matter. The *red contour* corresponds to the transformed gray-white matter boundary of the tissue block photograph. Jacobian determinants are in a plausible range, and the Jacobian maps are qualitatively similar.

physically plausible, and the structure of the Jacobian maps are qualitatively similar, indicating the reproducibility of the stage-1 registration across multiple stains.

The Ferritin, PLP, CD68 and pTDP43 images share the same staining defect at the top: in a narrow band only the hematoxylin counter stain is present because the

slides were mounted here when the primary and secondary antibodies were applied during the staining process. This artefact causes a slight but visible deflection in the registered contour, as the artefact represents a strong edge, and the contour is aligned to this edge in the relative absence of anatomical contrast. The deflection is clearly visible in the respective Jacobian maps, and completely absent in the case of the tincture stains that present with consistent anatomical contrast. In Figure 3.11 the only negative Jacobian determinant is observed at the intersection of the artefact with the outer edge of the histology section.

The tears of in the H&E image were largely ignored by stage 1, as there were no matching horizontal features in the tissue block that they could be aligned to. As a result, the estimation of the displacements in the torn region were predominantly based on the regularisation term, which was strongly predictive given the nearby sulcus and pial surface of the gyrus. As a general rule, tears or folds in the histology section should be segmented and masked in advance of the registration.

The variations of the Jacobian determinant observed in the white matter regions might be counter-intuitive, given the relative scarcity of anatomical features that could drive the registration in these areas. There are at least two reasons why we observed them:

1. There are many small features, such as enlarged perivascular spaces in the histology image, or vessel cross sections in the photograph, that MIND is sensitive to at the resolutions of the optimisation, therefore they drive the registration locally through the image dissimilarity term.

2. The alignment of strong features nearby necessitates large displacements, which influence the displacement vectors in these areas by coherence, which is enforced by the multi-resolution optimisation scheme and higher regularisation.

Based on the observation of both consistent and unique patterns in the Jacobian maps of the different stains, the observed variations are most likely a combination of both factors, and they could be suppressed with higher regularisation and an

additional pre-processing step that filters the tiniest features of the images (e.g. smoothing, mean shift filtering).

In summary, stage 1 proved to be relatively robust against dim gray-white matter contrast, but the registration of tincture-stained sections with consistent anatomical contrast are expected to provide more reliable results.

3.10 A comparison with ANTs SyN

In a final experiment, we compared the accuracy of the stage-1 pipeline to an established and widely used registration tool, ANTs [37]. Comparisons were made with both the Mattes mutual information and the cross-correlation metrics, that were used in a previous study [78] to register histology sections. The registrations were carried out on the same 14 CC and 14 HC sections that were previously registered to their corresponding tissue blocks with the stage-1 routine. For a fair comparison, all ANTs registrations were initialised using `antsAffineInitializer` with 30° increments over the whole circle. Various combinations of parameters were screened for each metric, starting from those that were recommended in the ANTs documentation. Empirically the following multi-resolution configurations were found to yield the best results with ANTs:

1. **ANTs SyN Mattes:** antsRegistration –dimensionality 2 –float 0 –output $outdir/ants.syn/moving_to_fixed –interpolation Linear –winsorize-image-intensities [0.005,0.995] –use-histogram-matching 1 -r $outdir/ants.syn/init.mat -m Mattes[$outdir/fixed.png, $outdir/moving.png,1, 20, Random, 0.2] -t affine[2.0] -c [1500 x 1500 x 1500 x 300 x 100 x 0, 1.e-7, 5] -s 5x4x3x2x1x0 -f 7x6x5x4x2x1 -m Mattes[$outdir/fixed.png, $outdir/moving.png, 1, 32] -t syn[0.25,3.0,1] -c [200 x 200 x 200 x 200 x 150 x 50, 0, 5] -s 5x4x3x2x1x0 -f 7x6x5x4x2x1

2. **ANTs SyN CC:** antsRegistration –dimensionality 2 –float 0 –output $outdir/ants.syn.cc/moving_to_fixed –interpolation Linear –winsorize-image-intensities

[0.005,0.995] –use-histogram-matching 1 -r $outdir/ants.syn.cc/init.mat -m Mattes[$outdir/fixed.png, $outdir/moving.png, 1, 20, Random, 0.2] -t

As the TIRL stage-1 routine uses binary masks for the registration, the masks were exported from the TIRL pipeline and the ANTs registrations were repeated with the masks. The previously generated contours of the tissue block photographs were transformed to histology space with the `antsApplyTransformsToPoints` tool (antsApplyTransformsToPoints –dimensionality 2 –precision 1 –input *blockpts* –*output*d/transformed_block_contour.csv -t [$d/moving_to_fixed0GenericAffine.mat, 1] -t $d/moving_to_fixed1InverseWarp.nii.gz), and the MCDs were calculated to measure the accuracy of the registrations in each case.

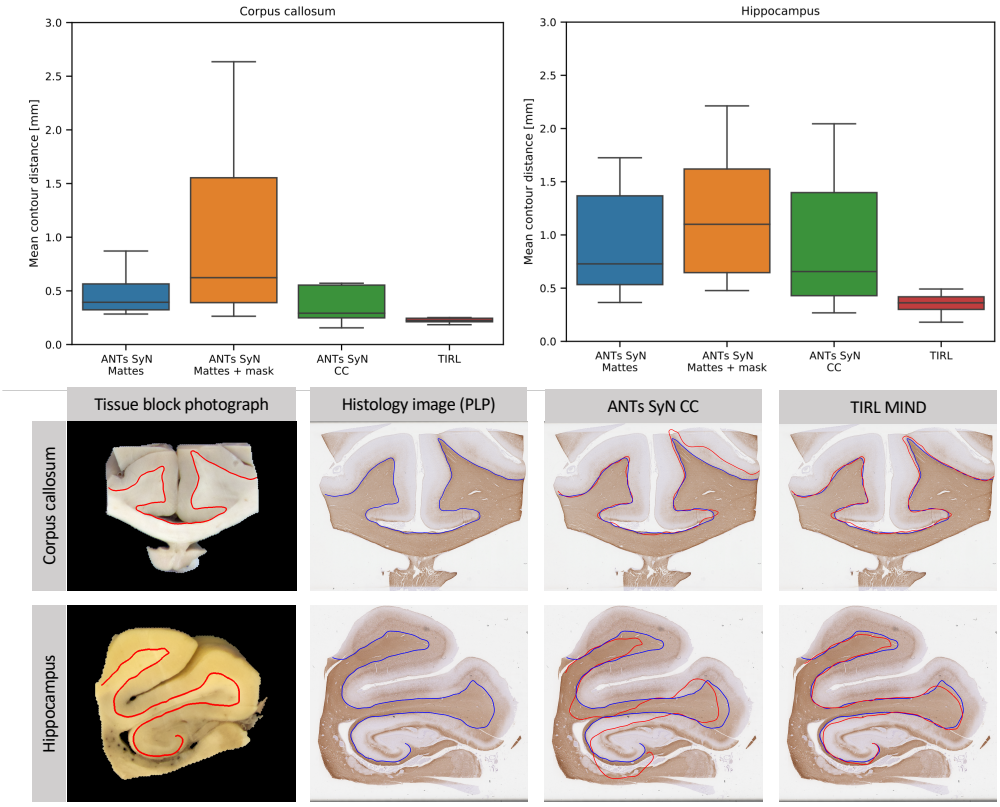

Figure 3.12: Comparison of stage-1 registration results with registration by ANTs. In the **top row** the results are pooled from all 14 callosal and 14 hippocampal slides. Cases where the affine initialisation failed were excluded from the ANTs results. In the **bottom two rows** are visual examples of representative registrations with ANTs SyN CC and TIRL MIND.

Figure 3.12 shows the pooled MCDs for each of the ANTs registrations compared to the TIRL stage-1 routine. Note that the masked registration is only shown for the Mattes metric, as the masks seemed to inflate the registration errors in both cases, and also led to several catastrophic misalignments in both cases. Both the hippocampal and the callosal sections could be more accurately registered with the CC metric than with Mattes metric, and the ANTs registrations were more accurate on the callosal sections than the hippocampal sections, similarly to the stage-1 routine. The most important result is that the stage-1 routine produced more accurate and more robust alignments, as can be seen in the boxplots. Representative registration results are also shown for comparison between the ANTs SyN CC and the TIRL results. The runtimes were comparable for ANTs and TIRL, averaging around 1 minute for ANTs, and around 1 minute and 15 seconds for TIRL.

The superior performance of the TIRL stage-1 routine can be attributed to the use of the MIND metric, which has superior sensitivity to image features, and a pixelwise optimised deformation field, which provided more flexibility to estimate local deformations.

4

Stage 2: Block-to-slice registration

Contents

4.1 Rationale 62
4.2 Image pre-processing 63
4.3 Automating block initialisation 65
4.4 Registration approach 67
4.5 Robustness of finding insertion sites 70
4.6 Initialisation of competing blocks 73
4.7 Results of block-to-slice registration 77
4.8 Effect of block initialisation on the registration accuracy 79

4.1 Rationale

The purpose of stage 2 is to align a photograph of a stand-alone tissue block to a photograph of the corresponding brain slice. To fully automate the reinsertion process, this stage uses additional photographs of the brain slice that help the localisation of the tissue block.

In the context of histology-to-MRI registration, stage 2 is an intermediate stage that alleviates the need for manually initialising a small 2-D histological section in 3-D MRI data. Not only is this a difficult and time-consuming task, that limits the throughput of the pipeline, but also it is error-prone. Certain tissue blocks exhibit unique anatomical features (e.g. the hippocampus) that make them easier to locate in MRI space. Others may only consist of a single gyrus or sulcus (e.g. primary motor cortex), that would be impossible to locate without prior anatomical information. In many cases, the initialisation has to be highly accurate even within the anatomical structure, as the image alone carries insufficient information for any registration method to adjust it to the exact 3-D position, 3-D orientation and 3-D deformations of the block – all of which would be necessary for a sub-voxel precise alignment. This requires careful practice from the operator, who needs to be looking at multiple slices of the MRI from various angles. Even with careful practice, human error should not be underestimated, especially in regions where consecutive MRI slices are highly similar (e.g. corpus callosum).

As an alternative to human intervention, the missing information can also be supplied by a brain slice photograph. The cross sectional anatomy (even of a single hemisphere) can provide sufficient guidance to determine the position and orientation of the slice in MRI space (this will be demonstrated in Chapter 5). If a brain slice can be accurately aligned with MRI data, it follows that a tissue block may also be accurately registered, provided that an additional 2-D registration is carried out that maps the points of the tissue block onto the respective ROI in the registered brain slice. This sets the objective for stage 2.

For stage 2, all input images are of the same modality and tissue deformations are also expected to be less prevalent than they were in stage 1, as formalin-fixed blocks

are fairly rigid. While these factors allow simplifications for the stage-2 approach, an important remaining challenge is to achieve reliable automated registration of the blocks that have less salient anatomical features, as these normally constitute a good fit at multiple regions within the brain slice. Accordingly, stage 2 consists of two steps: *block initialisation* and *block registration*. The two steps will be jointly referred to as *stage 2*, *block-to-slice registration* or *block insertion*.

4.2 Image pre-processing

4.2.1 Tissue block photographs

Tissue block photographs were already pre-processed for stage 1 (see Section 3.5), and the same images are used in stage 2.

4.2.2 Brain slice photographs

The original photographs of the brain slices were captured in the same autopsy sessions as the tissue blocks. Consequently, the size and resolution of the images were the same as for the tissue blocks ($\sim$3600 $\times$ 5400 pixels, $\sim$50 µm/px). The 10-15 mm thick coronal slices were placed on a blue matte background, excess formalin was wiped off with a tissue, and both the anterior and posterior surfaces were photographed. Small (approx. 25 $\times$ 30 mm) rectangular tissue blocks were sampled from various anatomical regions in each slice. Additional photographs were taken of both the anterior and posterior surfaces of the slice whenever a new block was removed to document the process of block sampling (on several occasions more than one block was removed at once). We distinguish the first image of each brain slice (where the slice is the most intact) by calling it the *brain slice photograph* or the *intact slice*, and the follow-up images that will be referred to as *auxiliary images* or *cut-out slices*.

All subsequent pre-processing steps were automated and run on the entire dataset (altogether 1073 images from 14 subjects[1]). The raw images were corrected for lens distortions and converted to TIFF format using the Darktable software [89].

[1]Two of the 16 subjects had to be discarded because of missing data.

Anterior photographs were flipped to enforce a consistent "neurological" orientation of all images. The images were cropped to remove auxiliary items on the periphery of the image. To identify the cropping region, the images were downsampled by a factor of 4 (∼0.2 µm/px) and converted into HSV colour space, and a binary mask was created from the hue channel by Otsu thresholding [93]. Objects were identified from the mask as connected components, and discarded from the mask if their centroid was outside a central rectangular region (50% of the image size). A bounding box was defined around the remaining objects (e.g. two hemispheres, or disjoint parts of a sampled brain slice) with an extra 10% margin in each direction. The original image was finally cropped to the upscaled bounding box, leaving only the brain slices and the blue background in the image.



Figure 4.1: Differences in the lighting necessitates different parameters for the segmentation of the brain slices.

The brain slices were segmented from the cropped images by pixelwise k-means classification. Background removal was implemented as a separate Python application with several options, mainly to accommodate varying lighting conditions (Figure 4.1) across different sessions (each brain was dissected in a separate session). Within the same session (individual), all images were segmented with the same configuration. Four different configurations were sufficient to produce good-quality segmentations for all images. These four configurations defined different feature vectors for classifying pixels. Feature vectors were created by a trial and error process from a combination of the most discriminative channels in the RGB, HSV, and L*a*b colour spaces: h (hue), sr (saturation+red), hsr (hue+saturation+red), ABsr

(a+b+saturation+red). $k = 2$ clusters were sufficient in all cases, except for one session where the $k = 3$ clusters were used with hsr vectors to accommodate specular reflections from the background sheath. Pixel-to-pixel variation in the images was suppressed by mean shift filtering [94] in RGB space (rgb=21, spatial=51px) before calculating the feature vectors. The component-wise demeaned and variance-normalised vectors were fed to the classifier to create a binary mask of the brain slice. Objects were identified in the mask as connected components, and discarded if they were smaller than 2000 pixels by area ($5\,\mathrm{mm}^2$) to exclude tissue fragments.

At the start of stage 2, the cropped and segmented photographs are converted to grayscale by extracting their luminance channel in YIQ colour space, as discussed in Section 3.5.

4.3 Automating block initialisation

The aim of block initialisation is to determine the approximate sampling location of one or more tissue blocks within a whole brain slice photograph. This is accomplished by comparing several photographs that show the brain slice at various stages of block sampling. In the absence of auxiliary brain slice photographs, this step of the pipeline can be skipped, and the location of the blocks in that instance should be defined manually. Figure 4.2 shows the complete stage-2 pipeline with a particular emphasis on the block initialisation steps.

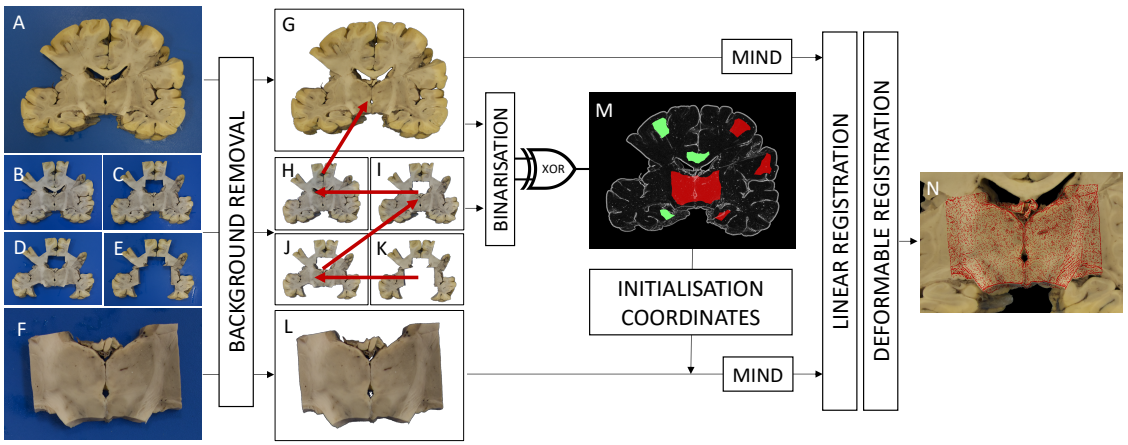

Figure 4.2: Automated initialisation and deformable registration of tissue blocks to the corresponding coronal brain slice photographs (see main text).

The brain slice photograph (A) and the corresponding cut-out images (B–E) are loaded, pre-processed and sorted according to the block sampling order, which is inferred from the number of brain pixels in the images. Consecutive images ($G \to K$) are then pairwise registered such that the less intact image (e.g. H) is warped onto the domain of the more intact image (e.g. G). Provided that the registrations are accurate, all non-overlapping regions should indicate sampling regions (M). The centre of each sampling region is determined and mapped to the intact brain slice photograph as a common reference frame. These coordinates will be referred to as *insertion sites*, and will serve as initial positions of the tissue blocks (F) when they are registered to the brain slice photograph (N).

The pairwise registrations are independent, hence they are carried out in parallel by TIRL. Each registration optimises a transformation chain that is attached to the domain of the more intact image. These transformation chains consist of a linear part with 5 DOF (rotation, anisotropic scaling, and translation) and a displacement field, which are optimised in succession. As the orientation of the brain slice images may vary, the linear registration commences with a rotation search at 1 mm resolution, measuring the normalised mutual information (NMI) of the images at 30° increments in the $[-180°, 180°)$ range. The best 3 rotation candidates are then selected and the rotation parameter is co-optimised with the scale factors and translation parameters using the BOBYQA method to maximise the NMI. The optimisation bounds are $\pm 15°$, $[0.8, 1.2]$, and ± 10 mm, respectively. The displacement field is defined on the image domain after the rigid registration, and optimised at 1.5 mm and 1.0 mm resolutions using diffusion registration (as described in stage 1). A slightly higher regularisation weight ($\alpha = 0.6$) was determined empirically to prevent nearby brain regions from expanding into the empty sampling regions. We noticed that the non-linear registration could not be fully eliminated, otherwise the genuine deformations of the brain slice would have caused misalignments that would have interfered with the detection of sampling regions.

To identify sampling regions, the registered images were binarised, and mutually exclusive regions were obtained by the XOR operation. Any region smaller than

1 cm^2 was discarded ("area filter"). The filtered binary image was converted to a Euclidean distance map, and values under 6 mm were discarded to exclude narrow bands around the edges of the brain slice that corresponded to misalignments that survived the area filtering ("streak filter"). All regions that passed the area and the streak filters were recognised as sampling regions, and the coordinates corresponding to the peak Euclidean distance were recorded. These points were mapped to the most intact brain slice photograph by inverting and concatenating the relevant slice-to-slice transformation chains. The transformed coordinates were eventually exported as insertion sites for the brain slice.

On certain slices the insertion sites were better identified if the number of sampling regions was manually specified for each pair of the brain slice images. In this case, the area and streak filtered Euclidean distance maps are binarised again, and the pixels are classified by their coordinates using k-means classification with the specified number of regions.

4.4 Registration approach

Block registration takes three inputs: (1) the cropped and segmented tissue block photograph, (2) the cropped and segmented brain slice photograph, and (3) a list of insertion sites. The images are converted to grayscale and the block is tested at all insertion sites. The algorithm determines the best site and refines the alignment of the block at this site by additional linear and non-linear registration steps.

Figure 4.3 shows the transformation chain that is optimised during block registration. Descriptions of the intermediate domains are given in Table 4.1. The chain (c_C) of the brain slice photograph (C) consists of a single transformation that scales the pixel coordinates of D_4 by the pixel size. The pixel size should be supplied by the user for each session by reading the visual scale in one of the autopsy photographs.

There is an equivalent initialisation part in the chain (c_B) of the tissue block (B), which scales the pixels by the resolution (*Resolution*), but this initialisation additionally moves the weighted centre of the image to the origin of S_1 (*Centre*). The

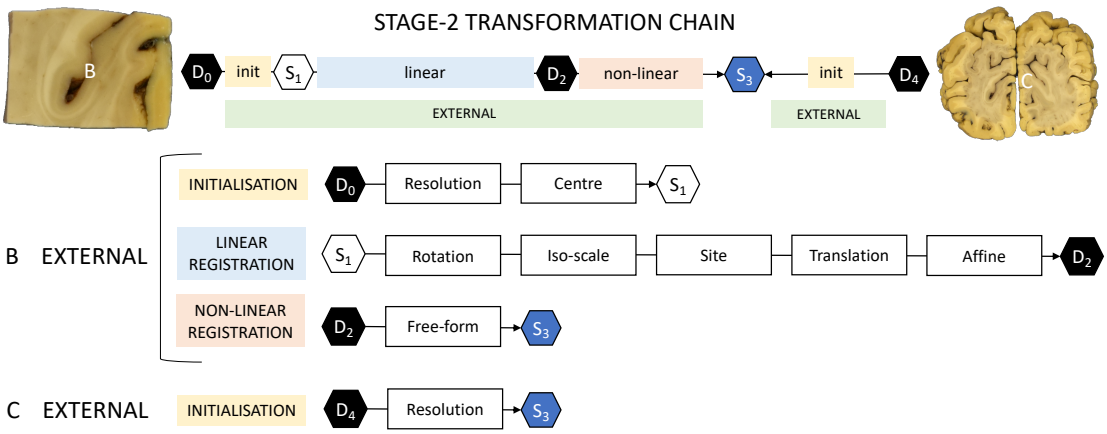

Figure 4.3: Diagram of the TIRL transformation chain that is optimised by stage 2 to align the tissue block photograph to the corresponding brain slice photograph.

Space	Description
D_0	the original pixel domain of B (50 µm/px)
S_1	B in metric space, origin at the weighted centre of the block
D_2	B in metric space at an insertion site after linear registration
S_3	B in metric space at an insertion site after linear + non-linear registration
	C in metric space (origin: top-left corner of C)
D_4	the original pixel domain of C (50 µm/px)

Table 4.1: Space and image domain definitions related to the stage-2 transformation chain.

linear subchain consists of a rotation, isotropic scaling, two translations (*Site* and *Translation*), and a full affine matrix. The *Site* transformation is set first to move the weighted centre of the tissue block photograph to one of the insertion sites, and the parameters of this transformation are kept constant every time until a new insertion site is tested. The non-linear subchain consists of a displacement field that is defined on D_2, and maps the points from the best affine alignment further towards S_3.

The dissimilarity of the images is quantified by resampling C on D_0 by the combined chain $(c_B)^{-1} \circ c_C$, and computing the MIND metric, as discussed in stage 1. Due to the part-to-whole nature of the registration task, it is inevitable that some of the tissue block background will overlap with brain regions in the brain slice. To suppress the effect of this apparent misalignment on the registration, MIND vector differences are only computed for the intersection of the brain areas by masking the input images. Masks are created automatically at the start of stage 2

by binarising the segmented input images at the 10% intensity level (relative to the maximum value).

The transformations of the c_C chain are optimised in 4 steps, as described in the following paragraphs.

1. **Finding the best insertion site.** To reduce computational costs, images are resampled to 0.5 mm resolution at the beginning of this step. The coordinates of a candidate insertion site are mapped to S_1, and assigned to the parameters of the *Site* transformation. At each site a grid search is performed on the rotation and translation parameters, such that the $[-180°, 180°)$ interval is tested by 30° increments, and the ±5 mm range is tested by 3 equispaced points. The MIND cost is evaluated at all combinations of parameters, and for the best 3 initialisations, the translation and rotation parameters are further refined by BOBYQA optimisation. Stringent parameter bounds are used to keep the optimisation close to the result of the grid search. The lowest of the three final MIND costs is recorded, and the same process is repeated at all insertion sites. Before testing a new site, all parameters are reset to their initial values. The best site is finally chosen on the basis of lowest overall MIND cost.

2. **Similarity transform.** The rotation angle, the isotropic scale factor and the two translation parameters are co-optimised in this step, starting from the final parameter values that were computed by the previous stage. The parameter bounds are expanded to accommodate a potential change in the scale factor: ±10°, 0.9–1.1, and ±5 mm, respectively. Optimisations are carried out at 0.5 mm and 0.25 mm resolution levels using the BOBYQA method.

3. **Affine transform.** The 6 parameters of the *Affine* transformation are optimised at 0.25 mm, 0.1 mm and 0.05 mm (native) resolution using the BOBYQA method. The parameter bounds are ±0.05 for the diagonal and off-diagonal elements and ±1.0 mm for the translation components, allowing small

adjustments to the previously optimised degrees of freedom while introducing shears.

4. **Diffusion registration.** This step optimises the vectors of the displacement field using the same method (Gauss–Newton) as described in stage 1. The field is optimised at 1.0 mm, 0.5 mm, and 0.25 mm resolutions, which we found to be a good compromise between accuracy and speed. The regularisation weight was empirically set to 0.6. Given that the registered images are of the same modality, and the deformations are generally small, the affine-registered images exhibit almost identical features after the MIND operation, and these are in almost complete overlap, hence a higher regularisation weight is appropriate to prevent spurious deformations at the finer scales.

4.5 Robustness of finding insertion sites

The robustness of the automated site finder routine was tested on the entire post-mortem dataset. This included coronal slices of all 14 subjects, that were sampled in at least one anatomical region (n=135). Photo series of the anterior and posterior surfaces of each slice were both included and treated independently, except for 12 slices where only the posterior surface was captured. The site finder routine was therefore run on 258 cases. The identified insertion sites were overlaid on the most intact slice, and visually compared with the least intact slice, looking for missing or misplaced points, that did not correspond to the tissue blocks.

In 190 cases (74%) the insertion sites were correctly identified for all blocks that were sampled from the slice. A typical example is shown in Figure 4.4. Among these, there were a few sporadic cases, where more insertion sites were found than necessary. These corresponded to tissue fragments that were inadvertently discarded during autopsy, especially where many blocks were removed from the slice.

In the other 68 cases (26%) only a subset of the insertion sites could be found with the default settings, and further information had to be provided to find all sites. Figure 4.5 shows typical examples where the algorithm did not produce

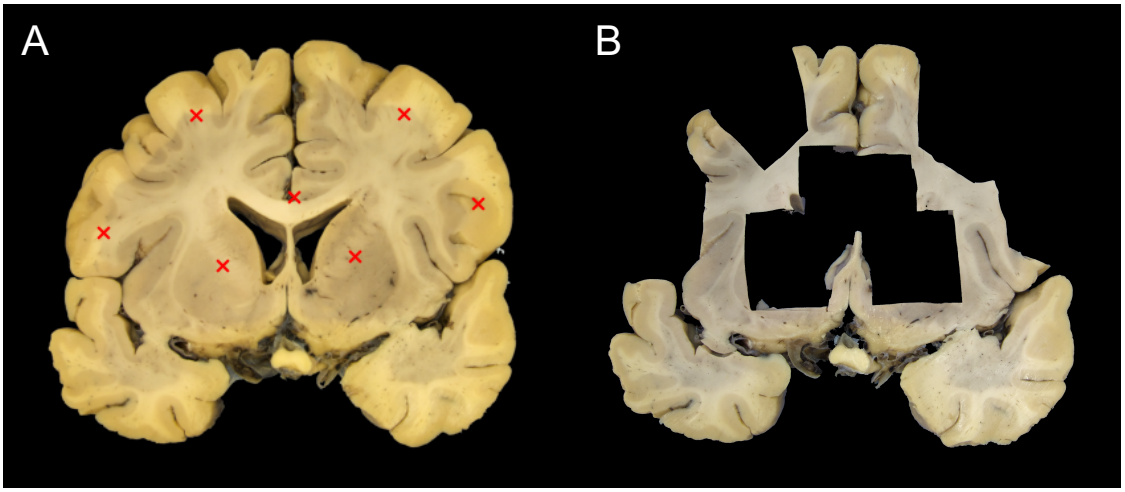

Figure 4.4: Representative example of a successful site finding process (7 sites). **(A)** Insertion sites (*red crosses*) overlaid on the most intact brain slice photograph. The sites are supposed to be close to the centre of the corresponding tissue block. **(B)** Least intact brain slice photograph, shown for comparison.

the expected result. An explanation of the cause and the applied correction is given below for each case.

A **Merged sampling regions.** The left and right orbitofrontal blocks were often detected as a single entity, because the blocks were removed at once, and the hemispheres were touching in this region. This was solved by manually specifying the number of blocks that the algorithm had to find based on each pair of input images. As the orbitofrontal blocks were sampled identically for all subjects (same slice number, one step), the script parameters only had to be modified once to obtain correct results for all subjects.

B **Split sampling regions.** A single sampling region is assigned to two or more insertion sites, which are therefore all offset from the true centre. This was observed in the majority of the callosal blocks that contained a segment of the corpus callosum and the cingulate gyri from both hemispheres. Depending on the interhemispheric distance, separate insertion sites were assigned to the cingulate gyri. Setting the number of expected regions to 1 in the respective block sampling stage yielded correct results in all cases. However, the exact parameters (which sampling stage, how many blocks per sampling stage) had

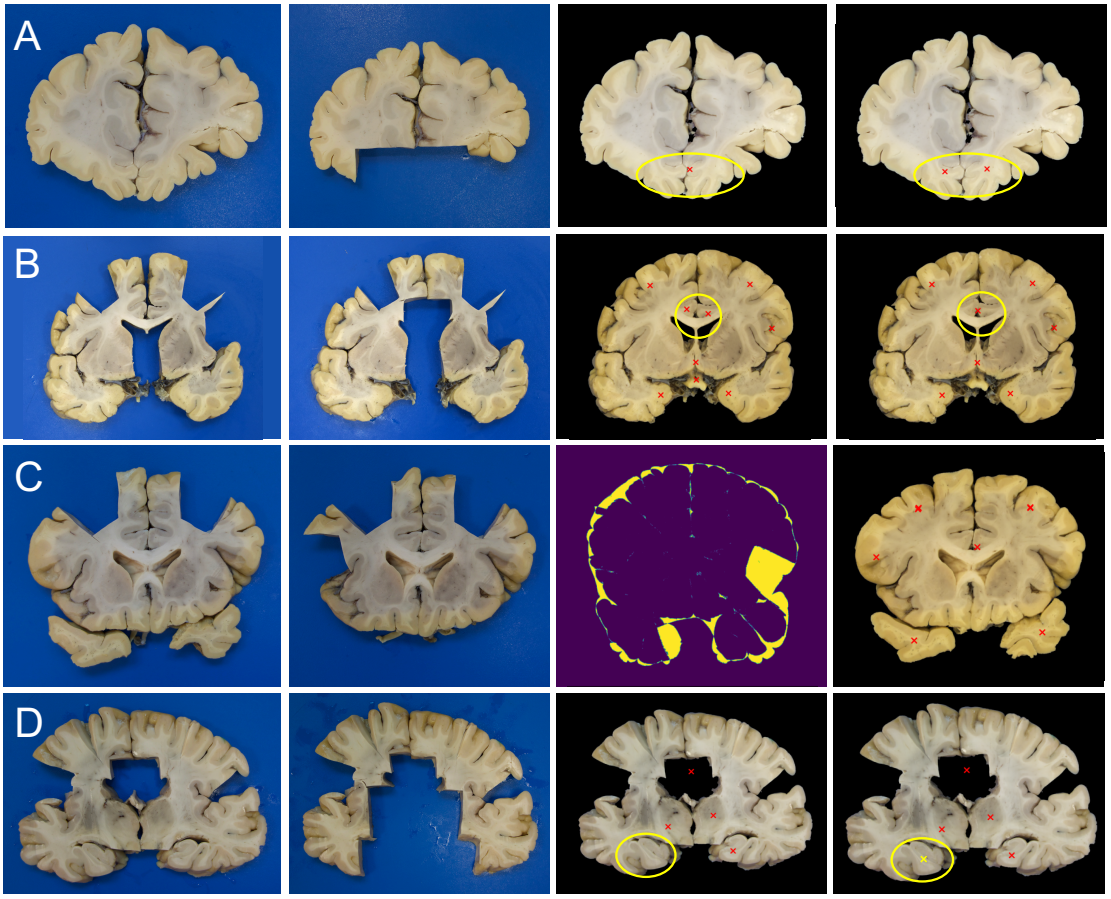

Figure 4.5: Examples where manual intervention was necessary to correctly identify all insertion sites for block initialisation. Each row shows a different type of problem (these are described in the main text). *Columns 1-2*: cut-out brain slice photographs that were directly responsible for the failure. *Column 3*: erroneous result (A,B,D: the incorrect site is highlighted by the *yellow ellipses*, C: erroneous orientation of one slice (*purple*) on another (*yellow*)). *Column 4*: most intact slice with the identified insertion sites after the correction.

to be tailored to the individual differences of sampling (slice number, number of fragments removed).

C **Complete dissociation of insertion sites from the sampling regions.** Sections that included the poles of the temporal lobe were often misrepresented in consecutive images, such that the temporal pieces appeared in the most intact state, but they were absent from subsequent images. As a result, the rotation search found an alternative orientation of the images that maximised the mutual information (see figure). Perhaps the MIND metric would have been more robust against this type of error, but skipping the rotation search

for these type of slices proved to be sufficient in all cases, as the images had upright orientation at the start.

D **Missing site.** Individual sampling regions were ignored for various reasons. The example shows a failed attempt of compensating for a missing intermediate image. As a result of human error, the hippocampi (2 blocks) and the large rectangular block with the thalami (1 block) were sampled all at once, with no intermediate photo taken. Setting the number of regions to 3 at this sampling stage resulted in one of the hippocampal sites being ignored, because the third ventricle had split the large block into two large regions, which were assigned to separate insertion sites, leaving only one space for the two hippocampal regions. Another reason for missed regions is when a misalignment of the slices splits the region into smaller ones that are ignored based on their small area. A third cause might be tissue debris that is connected to the region of interest if it is large enough to be mistaken for the actual region.

Among the above-mentioned cases, merging (24) and splitting (23) were the most prevalent, with relatively few cases of rotation (8). These could be compensated for semi-automatically as described above. Individual sites were missing in 6 cases due to slice-specific reasons - these were added manually. In the remaining 7 slices (3%) there were too many fragments to account for, and the insertion sites could be more conveniently defined by hand.

4.6 Initialisation of competing blocks

Typically more than one block needs to be inserted into a single slice (3-4 on average). Finding the appropriate insertion site for each block is crucial for the success of stage 2, as the subsequent block registration step is essentially confined to the region of the initialisation. Therefore the similarity metric that is used to choose the best site for each block must be able to discriminate between multiple insertions of the same block, and identify the best among these.

We compared the discriminatory power of the MIND similarity metric versus that of NMI in an extreme case where 8 blocks were sampled from the same coronal slice. Figure 4.6 shows the sampling regions (*white rectangles*) of the brain slice that correspond to the tissue blocks (*a–j*). The *red crosses* indicate the 10 automatically determined insertion sites (*A–J*), two of which are superfluous (*E, G*), as the corresponding tissue blocks were discarded during autopsy. The bar charts display the values of the MIND and NMI metrics of each block at each insertion site as calculated by the stage-2 block initialisation routine. All inputs and configurations were identical, except for the similarity metric, and NMI was calculated with 32 bins. Masked areas were ignored in both cases. Capital letters above the bars indicate the site with the largest NMI and the smallest MIND value for each block, which were therefore deemed best by the algorithm.

Based on the NMI metric, only 5/8 blocks were reinserted correctly, whereas all blocks could be reinserted correctly with MIND. Two of the failing blocks (*d*, *f*) were assigned to the false insertion sites (*E*, *G*). Had these two sites been excluded from the list, block *d* would have still registered to site *F*. This clearly indicates the challenge of discriminating between too many sites, some of which might mimic the correct site. Block *i* was inserted at site *B*, which may seem counter-intuitive as the blocks bear little resemblance on visual inspection.

The most likely explanation for the inferior performance of NMI in all of these cases is that NMI measures the similarity of the image histograms rather than the actual image features. Consequently, if a block or sampling region has a dominant subpopulation of pixel values, misalignments in other parts of the image may be less influential on the rotation and translation parameters, leading to a sub-par alignment of the block even at the correct initialisation site. Not finding the global optimum at the correct site leads to relatively uniform NMI values of the failing blocks at all sites, making the choice of the insertion site arbitrary in these cases.

As opposed to NMI, MIND detects local image features, and generates a large registration force when any of these are misaligned. This increases the chance of finding the global optimum even for the blocks that exhibit less salient anatomical

features (e.g. *a, c, d*). MIND values at the correct sites were found to be largely different from all other sites, for which it was relatively uniform. This shows a good discriminatory power even for as many as 10 sites, which rarely occurs in our dataset. The only exception was block *f*, where site *I* was a close competitor. This is due to the large separation of the hemispheres in block *f*, which caused a sub-par alignment with the sampling region, as the block initialisation routine only optimised the position and the orientation of the block, not accounting for any deformations.

The results of the experiment have indicated that MIND is superior to NMI for the identification of matching insertion sites, and it is possible to automatically register as many as 8 blocks to a single brain slice with the stage-2 pipeline.

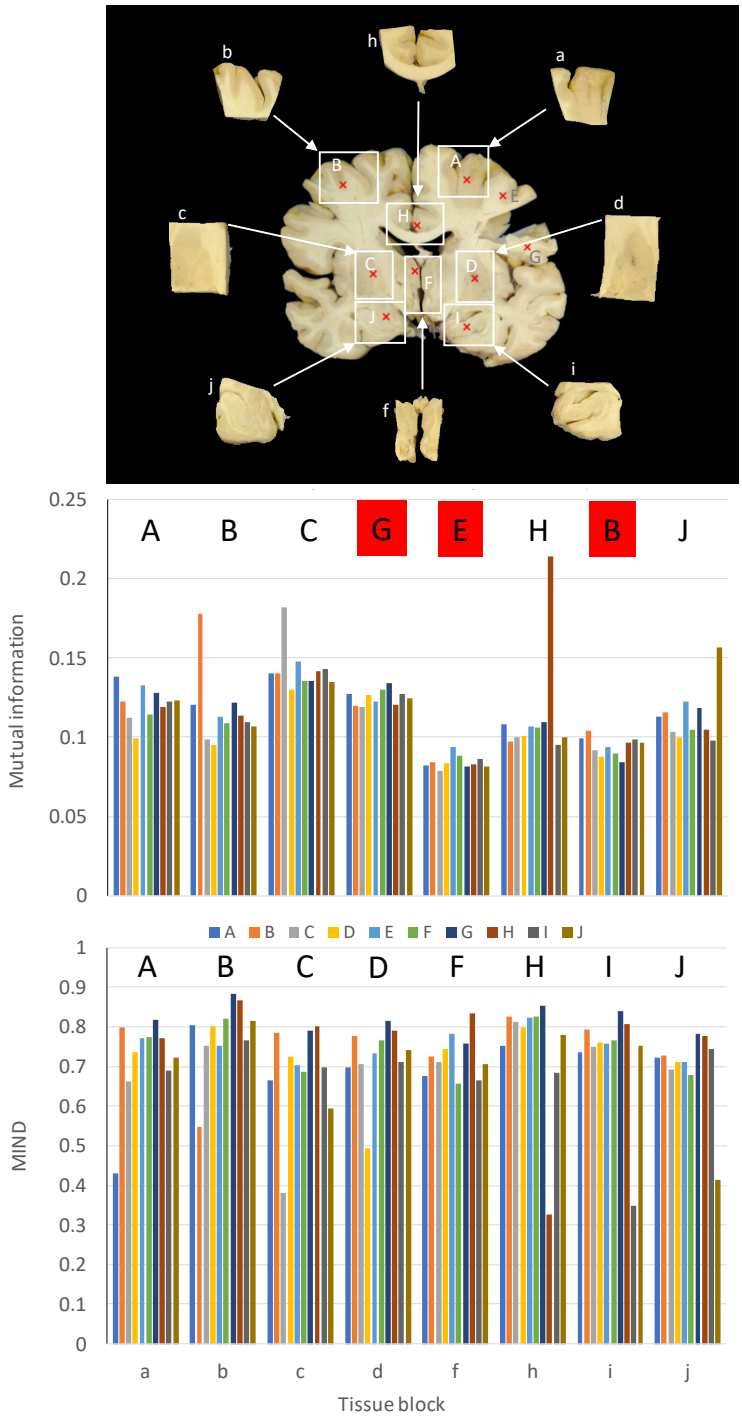

Figure 4.6: Top: sampling regions (A-J) with the respective tissue blocks (a-j). Note that *E* and *G* are false insertion sites with no corresponding block image. **Middle:** NMI values for all combinations of blocks (*categorical axis*) and sites (*coloured bars*). The capital letters above the chart show the region with the highest value of MI. *Red highlights* indicate where this leads to an incorrect matching. **Bottom:** The same matching process with the MIND metric. As opposed to NMI, smaller values of MIND indicate a better fit.

4.7 Results of block-to-slice registration

To evaluate the stage-2 pipeline as whole, we registered 87 tissue blocks from 14 subjects to the relevant brain slice images using identical configurations (Sections 4.3 and 4.4). These blocks are from the corpus callosum, anterior cingulate, hippocampus and visual cortex regions, and constitute the subset of blocks that have been processed for histology in the post-mortem ALS dataset prior to June 2020. Prior to the registration, insertion sites were identified for all slices using the automated routine. Missing insertion sites were manually defined because of missing intermediate cut-out images in 12 cases. The quality of the registrations were assessed by observation: the registered images were overlaid on top of each other, and the opacity was changed to transition from one image to the other while looking for any misaligned regions.

We observed that almost all registrations were highly accurate: gray-white boundaries were well-aligned, and characteristic small features of the images, such as penetrating vascular structures were generally within a 10 pixel distance ($< 0.5\,\text{mm}$) from each other. The perceived accuracy of the registration was highly consistent in all 81 cases, where the registration was deemed successful. In the remaining 6 cases we observed a failure, as the registered images beared little resemblance. With careful analysis, eventually all these failures could be traced back to human error, such as incorrect labelling of the slice photograph, the tissue block, misidentification of the corresponding slice, or block surface.

Figure 4.7 shows example registration results from various anatomical regions. These are representative of all 81 successful registrations.

To characterise the block initialisation error, we calculated the distance between the insertion sites and the centre of the inserted tissue blocks in all 87 cases. Figure 4.8 shows the distribution of initialisation errors with the median, the 75th, and the 95th percentile values.

At first sight, the low initialisation errors might be considered to be a result of a very restricted parameter range that the optimisation explores during the registration of the block. However, given that most registrations were accurate, the

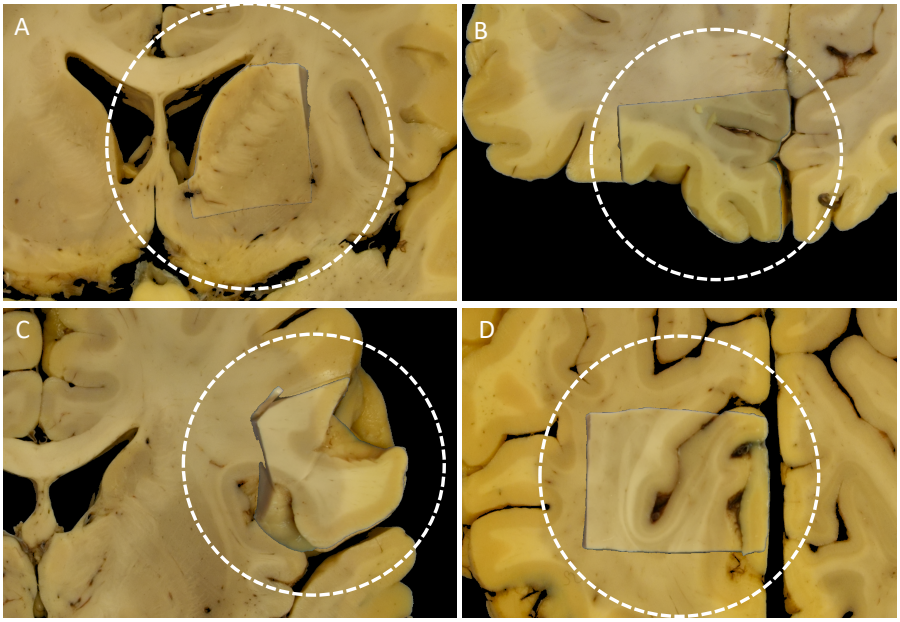

Figure 4.7: Example stage-2 registration results. **(A)** right internal capsule, **(B)** left orbitofrontal cortex, **(C)** Broca's homologue area, **(D)** left calcarine fissure.

small values are actually indicative of the highly precise automated initialisation in most of the tested cases.

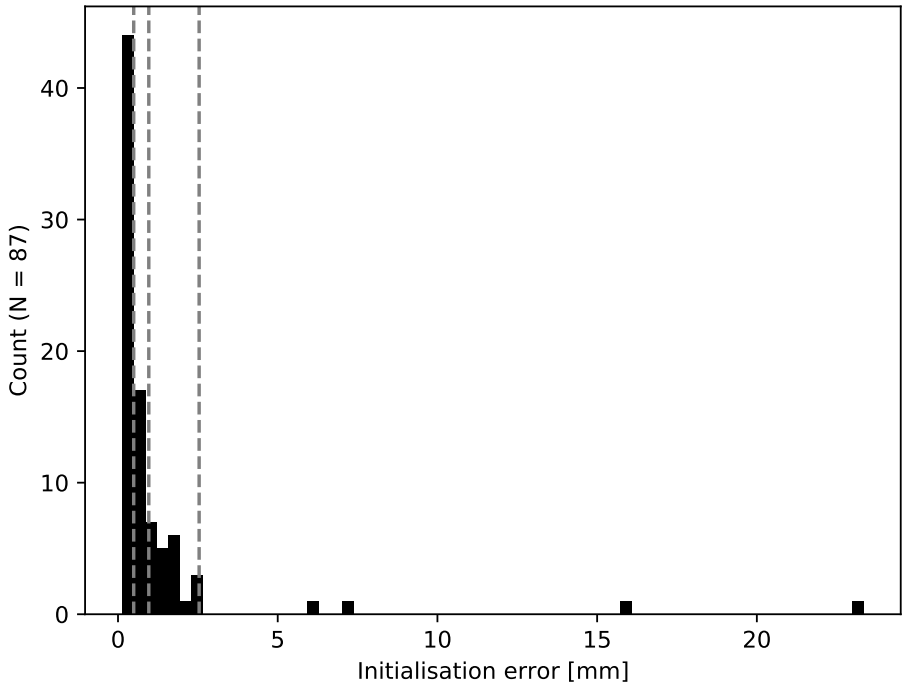

Figure 4.8: Error of block initialisation relative to the final position of the block, as measured on $N = 87$ blocks. The dashed lines indicate the median (0.49 mm), the 75th (0.96 mm), and the 95th percentile (2.54 mm) distances.

4.8 Effect of block initialisation on the registration accuracy

Finding the appropriate insertion site is the most important determinant of the success of stage 2. However, it is not clear how much the registration outcome depends on smaller inaccuracies of the initialisation. We therefore ran a final simulation experiment to test the robustness of the block registration step against variations in the initial position of the blocks.

Insertion sites were determined automatically for the same 8 tissue blocks that were used in the previous experiment, and a total of 100 initialisation points were sampled randomly around the original insertion site of each block from a Gaussian distribution with $\sigma = 10$ mm. The block registration routine was run separately from each point, using identical configurations. To assess the accuracy of the registrations, ground truth contours were created by hand around a salient anatomical feature in each of the tissue block photographs (as described in the stage 1 experiments) and matching contours were created in the respective sampling

regions of the brain slice photograph. The MCD was measured after the brain slice contours were transformed to the pixel domain of the tissue block. Any registration with $MCD \leq 0.2\,\text{mm}$ was deemed successful, anything above that was considered a failure. The choice of the threshold reflected the resolution at which the non-linear transformation was optimised.

Figure 4.9 summarises the results of the simulation experiment for all eight blocks. The sampling regions are shown with the true centre of the tissue block (*purple cross*), and the 100 initial positions scattered around the original insertion site. The true centre of the block is different from the insertion site, as the latter was inferred from the cut-out slices, and the former was calculated as the average weighted centre of the reinserted block from all successful registrations. Green dots represent the starting point of successful registrations, red dots indicate failures. The graphs show the robustness of the registration as a function of distance between the initialisation and the true centre of the block. More specifically, the *red curves* represent $100 - \#failures$ (primary y-axis), the *blue curves* represent the ratio of successful registrations versus all registrations within a given radius from the true centre (secondary y-axis). The dashed gray lines indicate the median and the 95th percentile of the automated initialisation errors as measured earlier on $N = 87$ blocks.

The majority of the registrations were successful for all blocks, up to the furthermost points (see the end of the red curve). As expected, the graphs show declining robustness of the registration for initialisations that are further away from the true centre of the block. No failures were observed within the median initialisation error for any of the blocks. Two failures were observed for blocks *a* and *d*, and one for block *j* within the 95th percentile radius. For blocks with less salient anatomical features (*a, c, d*) there was a marked decrease in the success rate past the 95th percentile error mark, which only occurred beyond 7.5 mm for the blocks that presented with clear contrast. In the case of block *f*, the original insertion site was far from the true centre, which explains the asymmetric spatial distribution of successes and failures. A similar asymmetry was observed in the case of block

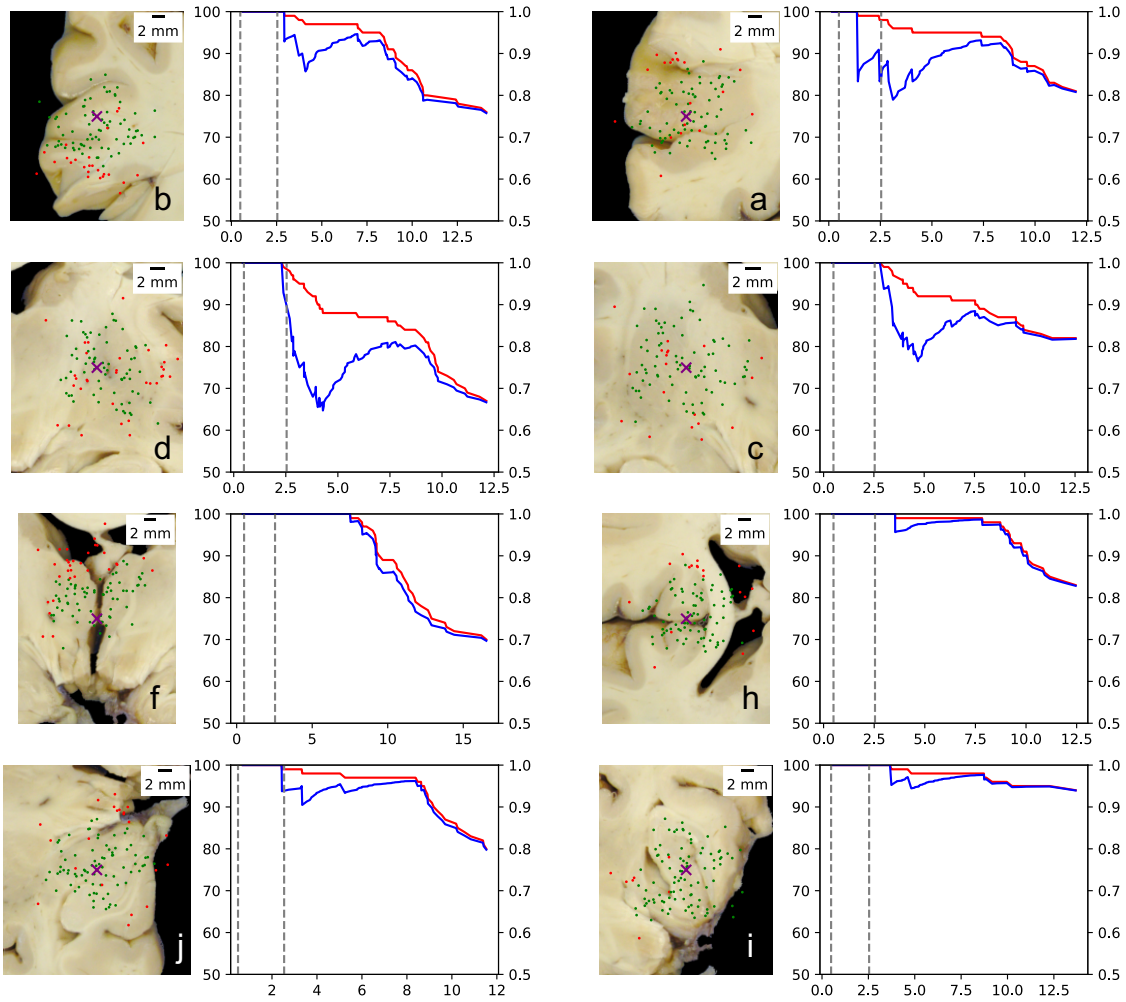

Figure 4.9: Registration summary after 100 random initialisations in the vicinity of the automatic initialisation points. The *purple cross* indicates the true centre of each block. *Green dots* indicate successful registrations, *red dots* indicate failures. The letters correspond to the block designations in the previous experiment. The charts represent success rates as a function of the initialisation error (see main text). The gray dashed lines represent the median and the 95th percentile automated initialisation errors.

b, where the spatially repeating pattern of gyri on one side started to attract the block at the most distant initialisations towards a different local optimum.

The results of the simulation experiment indicate that with the configurations outlined in Sections 4.3 and 4.4, stage-2 block registrations are highly accurate ($\leq$0.2 mm) in almost all cases when the block is initialised at automatically determined insertion sites. Failures can be more prevalent for blocks with less salient anatomical features, but only if the insertion site is offset by more than 2.5 mm. If a misregistration occurs as a consequence of erroneous initialisation, manual

intervention is needed to provide a more appropriate initial position for the blocks. Alternatively, the grid search of the rotation and translation parameters may be refined to increase the robustness of the pipeline at the expense of additional computational cost.

5

Stage 3: Slice-to-volume registration

Contents

5.1 Rationale 84
5.2 Image pre-processing 86
5.3 Transformation chain and brain slice initialisation . . . 88
5.4 Registration approach 92
5.5 Measuring the registration accuracy on simulated slices 97
5.6 Evaluating the MIND metric versus mutual information 101
5.7 Control point optimisation 104
5.8 Registration of a real brain slice photograph 114
5.9 Comparison with manual landmarks 116

5.1 Rationale

Stage 3 aims to register a brain slice photograph to 3-D structural MRI data. The points within the brain slice photograph are mapped to a surface in MRI space, that is not necessarily parallel to the MRI slices. Depending on the irregularities of free-hand cuts and the bulk deformations of the brain between the scanning and autopsy sessions, the target surface may also be curvilinear. In either case, finding the slice in the MR volume that corresponds to the slice in the photograph can be a difficult and time-consuming manual task. To achieve a precise alignment, the operator needs to optimise the viewing angle of the MRI about all three axes, where a slight deviation of 1° about a single axis is enough to offset cortical gray matter by one voxel (0.85 mm). This also means that larger initial rotations may cause difficulties recognising the slice, requiring many iterations. Figure 5.1 illustrates this problem by showing a comparison of the cross-sectional anatomy under various transformations of the slicing plane/surface. It is clear that a 5-mm offset or a rotation of 10° about a single axis is enough to cause some difficulty, but the combination of two rotations even without an offset makes the anatomy difficult to recognise as being from near the same location. In panel *E*, the added quadratic curvature of the slicing plane creates a persistent 0-5 mm offset in different regions that cannot be compensated by translations and rotations. Based on the details of the dissection protocol, finding one's way between oblique (let alone curved) representations of the anatomy can be a rather challenging task even for neuroanatomy experts. This is because they are generally more familiar looking at standard sections that often display anatomical structures symmetrically. The numbers that are used to illustrate the problem are based on observations in our post-mortem dataset, and are not meant to exaggerate the difficulty of this stage. It is necessary to automate this stage for higher throughput and increased precision in histology-to-MRI registration.

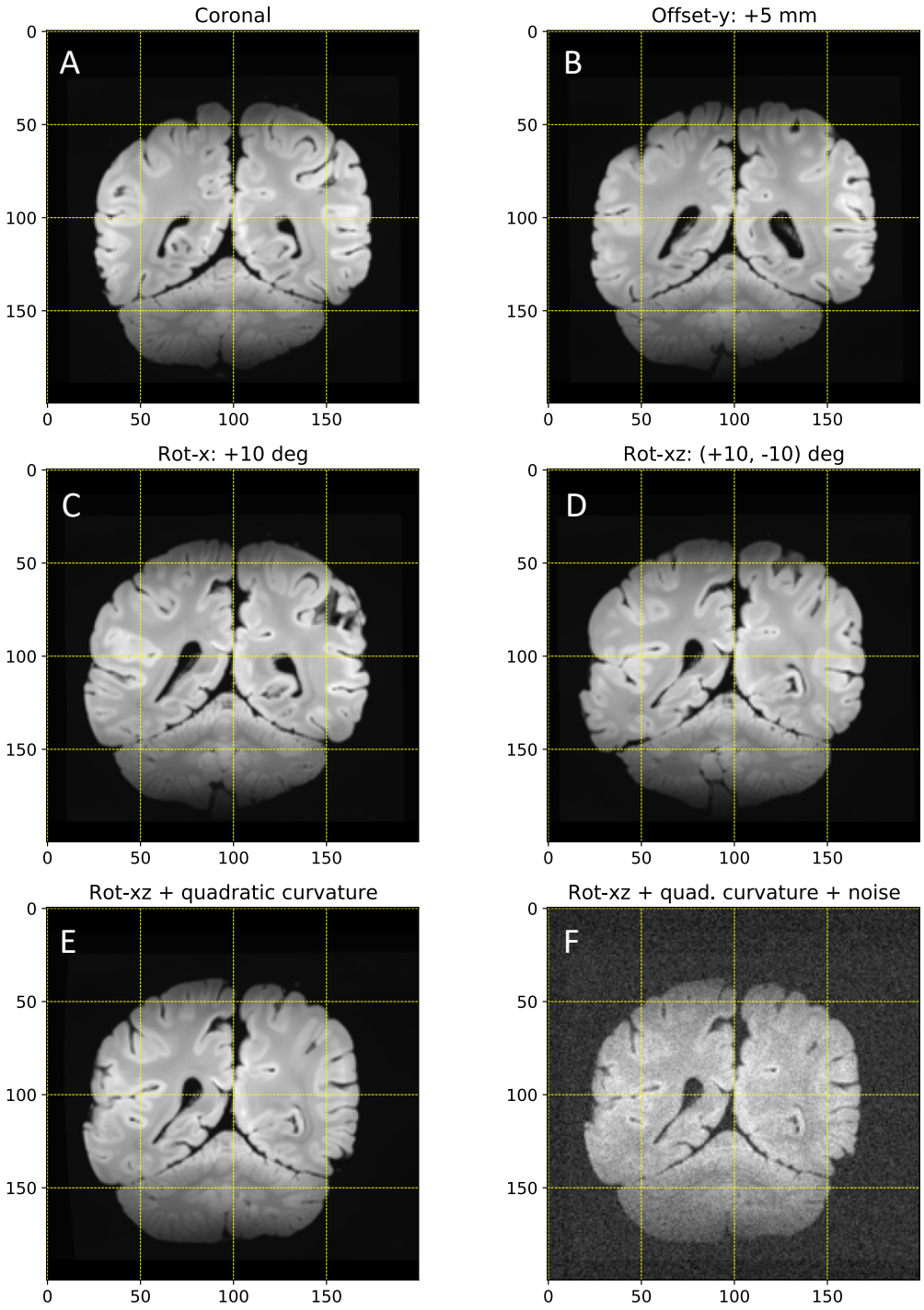

Figure 5.1: The difficulty of slice-to-volume matching is illustrated by comparing different cross-sections of the MRI volume. Despite the sections being relatively close, the cross-sectional anatomy varies largely between them. **(A)** coronal section, **(B)** coronal section shifted anterior by 5 mm, **(C)** coronal section rotated by 10° about the LR axis (x), **(D)** further $-10°$ rotation about the SI axis (z), **(E)** double-rotated slice with through-plane curvature. **(F)** shows (E) with Gaussian noise (SNR = 10) to imitate low gray-white contrast. *Yellow gridlines* are provided for comparison. (Resolution: 0.85 mm/vx)

5.2 Image pre-processing

5.2.1 Brain slice photograph

Stage 3 registers the pre-processed brain slice photograph from stage 2, which is cropped, neurologically-oriented, and background-segmented (see Section 4.2). To ease the computational cost, the images are converted to YIQ colour space, and only the luminance channel is used for the registration. For cost masking, a binary image is generated automatically by thresholding at the 10% intensity level (relative to the maximum intensity in the image). For images that contain visible side surfaces of the brain slice, the mask is edited manually to exclude these regions from the registration.

In the first step of the stage 3 pipeline, the images are downsampled to match the resolution of the MRI volume (0.85 mm/px). Since there is no unique information in the MRI data beyond this spatial resolution (as interpolation does not generate additional information, only measurement does), the associated computational cost of registering the images at higher resolutions can be saved.

5.2.2 MRI volume

Stage 3 uses the structural MR volume (TRUFI) that is resampled to the subject-specific common MRI domain, where the diffusion, susceptibility and relaxometry maps are also defined. This is a convenient choice for the target space for two reasons. Firstly, histology images will be registered to the common space, allowing direct comparison with multi-modal MRI data at each voxel. Secondly, at 0.85 mm resolution the runtime of stage 3 is only a couple of minutes, which makes it convenient to fine-tune stage-specific configurations for more accurate results. The best configurations at 0.85 mm normally represent a good starting point for registrations at higher resolution as well.

As the cerebellum is removed between the scanning and the dissection, additional pre-processing steps are necessary for each MR volume.

1. The missing cerebellar region is masked out from each volume by a binary cerebrum mask. These masks were generated automatically using the cortical surface extraction sequence of BrainSuite 19 [95]. BrainSuite was chosen because it was freely available, simple to use, and we found that the default configurations produced valid segmentations for the post-mortem images. Small inaccuracies were corrected by hand in two cases using ITK-SNAP [96]. The mask for each volume was specified as an input parameter for the registration.

2. Partly because of the missing cerebellum, the brain occupies a different resting position on the dissection table than inside the scanner. Consequently, the orientation of the physical cuts is most likely different from the coronal orientation of MR volume, and it helps the registration if the physical slicing orientation is approximated. For reasons that were outlined in the introduction of this chapter, this is a hard task, so we aimed for estimating the orientation of the middle slice in one brain only and only about the transverse axis to account for the tilt that is caused by the absence of the cerebellum. Using the Nudge tool of FSLeyes [97] the volume was rotated until the AP axis roughly coincided with the physical slicing orientation (i.e. the middle slice showed satisfactory resemblance with the relevant slice photograph). The rest of the volumes were then rigidly registered to this template by ANTs [37], and the unit AP vector was inverse-transformed to obtain the physical slicing orientation vector for each original volume. The physical slicing orientation was then specified as an input parameter for the registration for each volume. The main reason why ANTs was used for this task is that it automatically calculates the inverse transformations and it has a binary executable (`antsApplyTransformsToPoints`) for transforming arbitrary coordinate vectors by them.

3. Finally, MR signal intensities were normalised by the maximum value.

5.3 Transformation chain and brain slice initialisation

Figure 5.2 shows the transformation chains that are used by stage 3. As with the previous stages, there is a separate transformation chain associated with each of the input images. These converge on the S_4 metric space (physical MRI space), where the brain slice photograph is represented on a close-to-coronal 3-D surface domain. Table 5.1 summarises notable intermediate spaces of the registration.

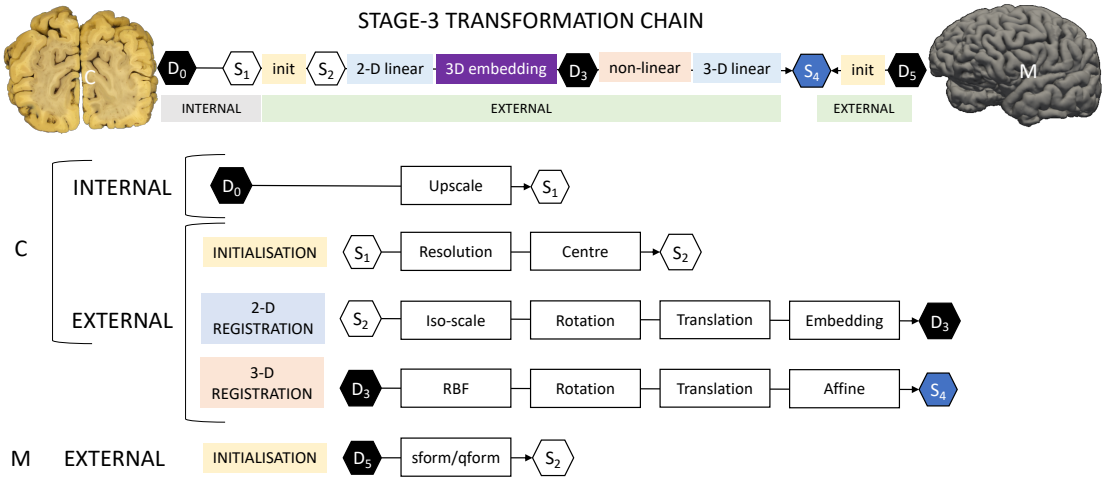

Figure 5.2: Diagram of the TIRL transformation chain that is optimised by stage 3 to align the brain slice photograph with volumetric MRI data. C and M refer to the brain slice photograph and the MRI volume as image functions, respectively.

Space	Description
D_0	downsampled pixel domain of C to match the MRI resolution (0.85 mm/px)
S_1	original pixel domain of C (50 µm/px)
S_2	C in 2-D metric (mm) space, origin at the weighted image centre
D_3	C in 3-D metric (mm) space after in-plane linear registration
S_4	C in 3-D metric (mm) space after all registration
	M in 3-D metric (mm) space after initialization by *sform*
D_5	voxel domain of M (0.85 mm/vx)

Table 5.1: Space and image domain definitions related to the stage-3 transformation chain.

The MR volume has a single affine transformation (equivalent to the NIfTI *sform* matrix [29]) that maps its voxels into a physical reference frame. This is a RAS

coordinate system with millimetre units, and the origin is at the centre of the brain, which is calculated as the mean of the physical coordinates of all non-zero voxels).

The transformation chain of the brain slice photograph is quite complex, including both an internal and an external part. The internal part is generated automatically by TIRL when the brain slice photograph is downsampled (this first happens in the pre-processing, and several times afterwards as part of the multi-level registration scheme). The internal part scales the pixels of the downsampled image (D_0) back to the pixel domain of the original photograph (S_1). The first two transformations of the external part *initialise* the brain slice photograph in 2-D metric space (S_2), where coordinates are expressed in millimetres and the origin is at the weighted centre of the brain slice photograph. In later parts of the transformation chain, starting from S_2, the brain slice is first transformed by a chain of 2-D linear transformations (*blue*: isotropic scaling, rotation and translation), then it is embedded (*purple*) in the $z = 0$ plane in 3-D space (D_3), and transformed both linearly (*blue*) and non-linearly (*beige*) in 3-D. This organisation allows separate estimation of the in-plane and through-plane deformations of the slice relative to the MR volume.

The type of the non-linear transformation object *RBF* is different from the displacement field type (`TxDisplacementField`) that was used in stages 1 and 2. Displacement vectors for this object (`TxRbfDisplacementField`) are defined on a *sparse domain* (`Domain`), which consists of a small number (typically $8 \leq n \leq 64$) of *control points* (nodes) that are scattered across D_3:

$$\mathbf{x}_i \in D_3 \quad \text{for} \quad i = 1..n \tag{5.1}$$

Based on the sparse definition of n 3-D displacement vectors at n control points ($\mathbf{u}^{(i)}$), a dense 3-D displacement field $\mathbf{u}(\mathbf{x}) = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^\top$ is computed for $\forall \mathbf{x} \in D_3$ as a weighted combination of n Gaussian basis functions (G):

$$u_j(\mathbf{x}) = \sum_{i=1}^{n} w_{ij} G(r_i, \sigma) \tag{5.2}$$

where $r_i = \|\mathbf{x} - \mathbf{x}_i\|_2$, and σ is equal to the average pairwise distance of the control points. As f vanishes with increasing r, the support of the displacement

vectors in the dense field is predominantly local, and consists of the nearest few control points. The node weights (w_{ij}) for each node i and vector component j are obtained by solving the following linear system using the known displacement vectors and pairwise control point distances (r_{ik}):

$$\begin{bmatrix} f(r_{11}) & f(r_{12}) & \dots & f(r_{1n}) \\ f(r_{21}) & f(r_{22}) & \dots & f(r_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ f(r_{n1}) & f(r_{n2}) & \dots & f(r_{nn}) \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ \vdots & \vdots & \vdots \\ w_{n1} & w_{n2} & w_{n3} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ \vdots & \vdots & \vdots \\ u_{n1} & u_{n2} & u_{n3} \end{bmatrix} \tag{5.3}$$

The dense displacement field is thereafter used to map the 3-D coordinates of the newly embedded brain slice photograph ($\mathbf{x} \in D_3$): $\mathbf{x}' = \mathbf{x} + \mathbf{u}(\mathbf{x})$. This is followed by the linear transformations (3-D rotation, 3-D translation and 3-D affine) that map $\mathbf{x}'$ onto S_4. The order of the transformations ensures that the xy component of the displacement is always tangential, and the z component is always orthogonal to the slice, regardless of the orientation of the slice in 3-D space. The control points of the non-linear transformation can be defined either automatically or manually. Different automated control point placements will be compared in Section 5.7. Although manual control point placement is hardly ever required, to make it as convenient as possible, control points are defined at the native resolution of the photograph (in S_1) and are automatically mapped to D_3 (3-D) by TIRL. This also ensures that the non-linear transformation moves with the image even when the 2-D part of the chain is changed.

The parameters of the 3-D *Rotation* object are extrinsic *Euler*-angles about the z (IS), y (PA), and x (LR) axes of the physical MRI space (S_4). Since the brain slice photographs are coronal, the x-rotation is initially set to 90°. Furthermore, the 2-D rotation parameter is set to an initial value of −90°, because the x-y axes of D_0, S_1, and S_2 correspond to the SI and LR directions of the slice. Figure 5.3 shows the full initialisation sequence of the brain slice in MRI space before the registration.

At the end of this automatic initialisation, the slice is in the coronal plane, at the centre of the brain. Any further refinement of the slice position or the orientation must come from user input. As the exact position and orientation of a slice are often too hard to determine, the user is merely expected to provide a plausible

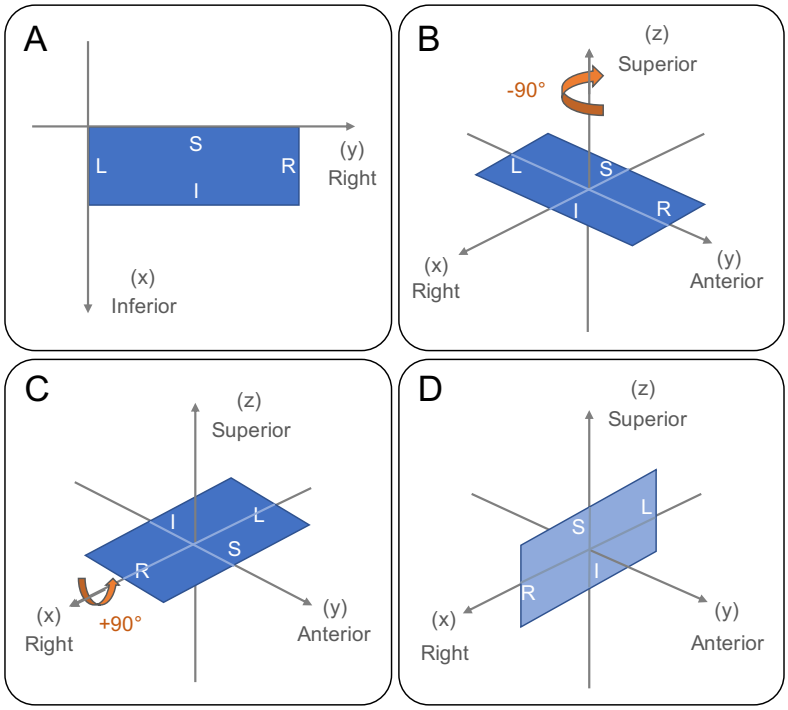

Figure 5.3: Schematic illustration of the brain slice initialisation sequence in physical MRI space, tailored for coronal slice orientation. Alternative initialisation sequences can be defined by changing the initial parameters of the stage-3 transformations. S: superior, I: inferior, L: left, R: right. **(A)** Brain slice in pixel space (D_0), **(B)** $-90°$ in-plane rotation after embedding at $z = 0$, **(C)** 90° rotation into the coronal plane, **(D)** Initialised brain slice in S_4 at the origin of MRI space. The position and orientation (see main text). The *lighter blue* colour indicates that the slice is now shown in radiological orientation. (Note that "surfaces" are degenerate for 2-D image data.)

range for the parameters of the 3-D translation and 3-D rotation components, and further refinement will be carried out by TIRL during the registration (this will be discussed in detail in the next section). A typical user input would be a 2 cm wide range along the sagittal axis for the position, and the sagittal axis for the slice orientation. To be more precise with the orientation, the physical slicing orientation can be estimated as described in the pre-processing section (Section 5.2).

TIRL has an additional utility (`multi_stage3_prepare`) to initialise a batch of parallel slices from the initialisation of the central slice. Since all slices are parallel, the same orientation estimate is specified for all slices, but the position estimates are shifted along the slice normal according to the slice spacing and the slice number. Further explanation on the use of this tool will be given in Chapter 6.

5.4 Registration approach

The 2-D linear, 3-D linear, and 3-D non-linear transformations of the chain are optimised in four substages, as illustrated by Figure 5.4.

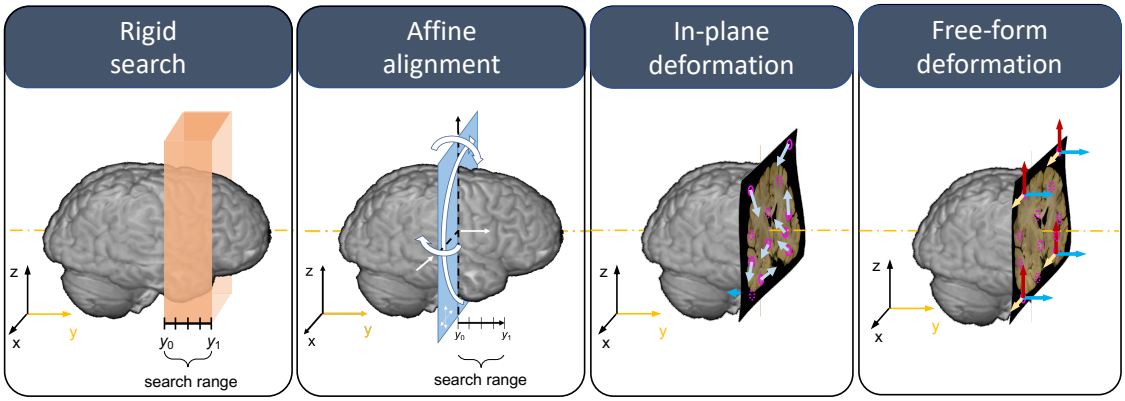

Figure 5.4: Substage overview of the slice-to-volume registration performed by stage 3.

5.4.1 Substage 1: rigid search

Stage 3 relies on a robust rigid initialisation of the slice. The 3-D position and 3-D orientation of the slice are therefore fine-tuned within a rectangular search region (slab) at the start of the registration by grid search. The slab is centred at the initial position of the slice, and oriented according to the pre-specified physical slicing orientation. Based on our initial experiments, the slab thickness should not exceed 20 mm. The position parameters are tested in discrete steps along the centre line of the slab, such that the distance between consecutive steps is equal to the voxel size, or it is on the order of a few millimetres (e.g. 2.5 mm). At each position, the *Euler*-angles are fine-tuned in a narrow range (e.g. $[-20°, 20°]$) in typically 3 or 5 steps. A full grid search with 11 positions and 5 steps for each rotation parameter requires $11 \times 5^3 = 1375$ independent tests, which are carried out in parallel by TIRL. Optionally, a series of N grid searches can be conducted with iteratively refined grids. On the second and further iterations the search region is gradually confined by re-centering the slab at the previous best position and allowing it to span between the neighbouring test positions only. Using $N > 2$ is rarely justified

by the accuracy of the final registration. Nevertheless it may be necessary if the initial test positions are sparse (e.g. more than 5 mm apart).

The linear transformation parameters (except for the affine) are locally optimised for each combination of initial translation and rotation parameters. During the optimisation, first the 2-D rotation and 2-D translation parameters are co-optimised with the 2-D isotropic scale factor, and this is followed by a joint optimisation of the 3-D rotation and 3-D translation parameters with the 2-D isotropic scale factor. Both the 2-D and the 3-D optimisations employ the BOBYQA algorithm and minimise the MIND cost between the brain slice and the MRI data that is resampled on the S_4 surface after each parameter change. All optimisations are carried out at 2x downsampled resolution and only the best 30% of the initialisations are fine-tuned at native MRI resolution, to reduce the computational load. Importantly, the parameter bounds are dynamically updated after the 2-D and 3-D optimisations by a pre-specified update rule (stage-3 configurations). The update rule allows for large excursions of the slice in the slicing plane (± 10 mm and $\pm 20°$), and moderate changes of the scale ($\pm 5\%$). The bounds for the 3-D translation parameters and the other two *Euler*-angles are equal to half of the step size that was used in the grid search for these parameters, so that they remain close to their initial values.

At the end of this substage, the parameters associated with the lowest MIND cost are assigned, and kept constant throughout the subsequent substages of the registration. Without further optimisation, the S_4 domain corresponds to the best rigid alignment of the brain slice photograph with the MR data.

5.4.2 Substage 2: affine alignment

This substage starts from the best rigid alignment of the slice, and optimises the 12 elements of the 3-D affine transformation matrix. The optimisation is carried out at native MRI resolution, and uses the BOBYQA algorithm to minimise the MIND cost. The parameter bounds are ± 1 mm for the translation components and ± 0.1 for the rest of the matrix components. Without further optimisation, the S_4 domain represents the best affine alignment of the slice with the MR data.

5.4.3 Substage 3: in-plane deformation

Following the affine step, any misalignment between the brain slice photograph and the resampled MRI data is due to non-linear deformations of the slice. Some regions may be deformed directly by the pressure that is exerted by the cutting, others may deform because they are no longer compressed under the weight of other parts of the brain. These effects are confined to the slicing surface, and an attempt is made to compensate for these at this substage by non-linear registration.

The *RBF* transformation object is initially created with n (usually 4) control points, and placed at its designated position in the transformation chain. Substage 3 then finds the optimum xy component of the n displacement vectors ($2n$ parameters) by minimising the MIND cost using the BOBYQA algorithm. To ensure smooth deformations, the MIND cost may be optionally supplemented by a regularisation term that penalises the second derivatives of the dense displacement field:

$$\sum_{\mathbf{x}\in D_3} \left\| \nabla^2 \mathbf{u}(\mathbf{x}) \right\|_F \tag{5.4}$$

where $\|\cdot\|_F$ is used to denote the Frobenius-norm, and ∇ is the vector derivative. However, in most cases the regularisation is not necessary, as the typical optimisation step, parameter bounds and number of control points are all small enough (0.5 mm, $\pm$5 mm, 16 points) to yield a smoothly varying deformation field with relatively small displacements. Hence by default, regularisation is disabled in the stage-3 routine. For closely spaced control points and large deformations, a regularisation term may be necessary.

To increase the chance of finding the global optimum, the control points are recruited gradually. After the first optimisation, a new *RBF* object is generated with $2n$ points ($4n$ parameters), and the new xy components are initialised to recreate the previously computed displacement vectors at all points of D_3. This requires solving an overdetermined linear system for the node weights, which is similar to Equation 5.3, but for $\forall \mathbf{x} \in D_3$ instead of n points. As the new field has twice as many degrees of freedom, the second optimisation fine-tunes the previous field estimate for more local, smaller-scale deformations of the slice. The optimisation

then continues for $3n$ and $4n$ points accordingly. Typical values for n are 4, 6, 8, 12, and 16. As the MIND cost depends on the scale, the four-step optimisation is carried out first at 2x downsampling and the final field estimate with all $4n$ points is later refined at native MRI resolution using the same $4n$ points.

At the end of this substage the S_4 domain represents the best non-linear estimate of a *planar* brain slice.

5.4.4 Substage 4: free-form deformation

Any remaining misalignment after substage 3 is most likely a result of the bulk deformations of the brain during dissection, part of which is orthogonal to the cutting surface. Another source might be the unevenness of free-hand cuts, leading to a slightly curved cutting surface in each case. These necessitate the introduction of additional degrees of freedom to deform the slice in the orthogonal direction.

The implementation of substage 4 is identical to that of the previous substage. The only difference is that substage 4 optimises all three components of the displacement vectors of the *RBF* transformation, starting from the final xy estimates of substage 3. This ensures that spurious through-plane deformations are not introduced to compensate for misalignments that are purely in-plane. At the same time, xy and z components are co-optimised, allowing correction of any deformations that were previously incorrectly attributed to in-plane causes, and trade them for the correct through-plane deformations. The initial step and parameter bounds for the z components should be determined empirically. Typical values are 1.5 mm and 5 mm, respectively. The purpose of the regularisation in substage 4 is to prevent a sharp bending of the slice. However, if the control points are well scattered and the curvature of the slice is small (< 5 mm), this is generally not required.

At the end of this substage the S_4 domain represents the best non-linear estimate of a *curvilinear* brain slice.

5.4.5 Substage 5: affine re-estimation

The linear substages may sometimes converge on a suboptimal rigid/affine alignment. Nevertheless, if the parameters are sufficiently close to the global optimum, the orthogonal deformations of substage 4 may compensate for this, incurring an affine component to the displacement field. Substage 5 therefore takes the stringent assumption that all through-plane deformations are artificial, arising from suboptimal affine initialisation. Consequently, substage 5 re-estimates the 3-D affine parameters based on the displacement field, and discards all through-plane deformations.

The following linear system is solved to obtain the least-squares estimate of the incremental affine matrix ($\mathbf{A}$):

$$\mathbf{AX} = \mathbf{Y} \tag{5.5}$$

where $\mathbf{X}$ is a $3 \times m$ matrix representing the physical coordinates of the linearly registered brain slice, and $\mathbf{Y}$ is the equivalent $3 \times m$ matrix representing the same coordinates after applying the orthogonal deformations. Here the xy non-linear components are ignored, as these have no influence on the slice position and orientation, but may adversely affect the above estimation, which depends on point-to-point correspondences. Furthermore, the matrices $\mathbf{X}$ and $\mathbf{Y}$ only contain coordinates from the brain mask, to exclude large local deformations close to the edges of the slice, which can greatly offset the estimation, despite having little impact on the actual alignment.

The new affine parameters are obtained by left-multiplying the old affine by the incremental affine. At the end of substage 5, the S_4 domain represents the best affine alignment of the slice with in-plane deformations. Stages 3-5 may also be repeated several times to re-estimate in-plane deformations from the new affine initialisation, and gain more confidence that the orthogonal deformations of the slice correspond to actual curvatures of the cutting surface.

5.5 Measuring the registration accuracy on simulated slices

Measuring the accuracy of the stage-3 protocol necessitates the annotation of anatomically correspondent points, curves or regions for the MRI data and the registered brain slice photographs. However, these are particularly challenging to generate due to the 2D/3D aspect (as discussed earlier in Section 5.1), and the errors of the annotation would also inflate the registration error. For these reasons, instead of using the slice photographs, we tested the accuracy and resilience of the stage-3 protocol on simulated slices, that were generated from the MRI data to mimic slice photographs. With the simulated slices, we aimed to test the performance of the automated slice-to-volume registration under the same conditions that were used to illustrate the difficulties of manually finding a slice in MRI space (Figure 5.1).

Figure 5.5 shows the four sampling schemes that were used in the experiment:

1. **Straight**. Planar sections perpendicular to the AP axis.
2. **Oblique**. Planar sections rotated by 10° about the AP and IS axes.
3. **Straight slices with curvature**.
4. **Oblique slices with curvature**.

The slices were created by resampling the 3-D MR data on a 200×200 pixel (17×17 cm) large *sampling frame* (`Domain`). The sampling frame was initialised by the default stage-3 transformation chain, and the chain parameters were adjusted according to the specific sampling scheme to generate 10 slices with 15.4 mm spacing for each series. For the planar slices only the 3-D translation and 3-D rotations were adjusted. To create curvilinear slices, the orthogonal displacements of the *RBF* transformation were parametrised with a 2nd-degree polynomial:

$$u_3(\hat{\mathbf{x}}) = c_1(\hat{x} - \hat{x}_0)^2 + c_2(\hat{x} - \hat{x}_0) + c_3(\hat{y} - \hat{y}_0) + c_4(\hat{y} - \hat{y}_0)^2 \tag{5.6}$$

where $\hat{\mathbf{x}}$ are the pixel coordinates of the sampling frame: $\hat{\mathbf{x}} = \begin{bmatrix} \hat{x} & \hat{y} \end{bmatrix}^\top$, and $(\hat{x}_0, \hat{y}_0)$ is the centre of the sampling frame. The coefficients were randomly sampled from

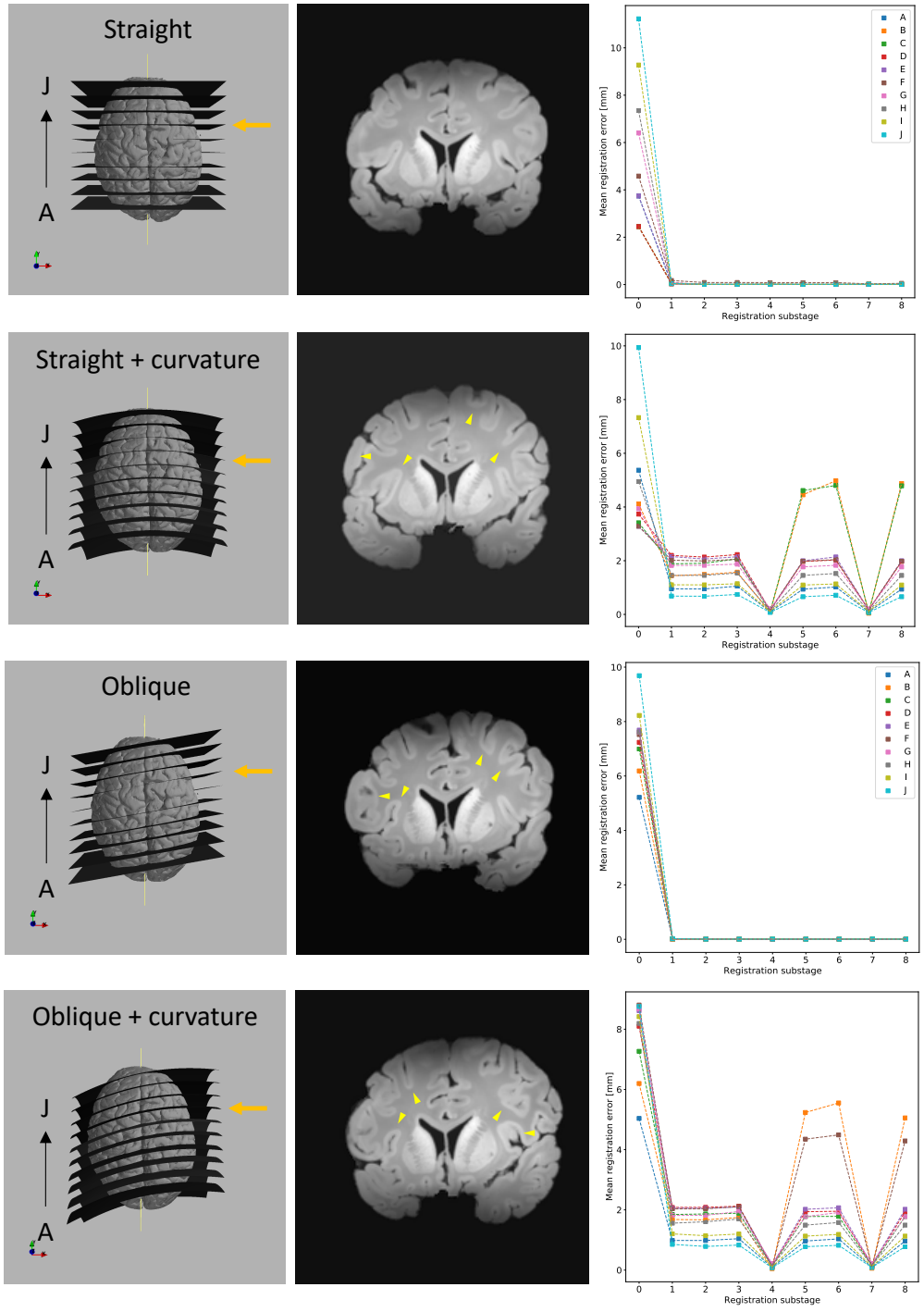

Figure 5.5: Slice-to-volume registration experiment with four series of simulated slices, corresponding to gradually increasing registration difficulty. The slices are labelled *A*–*J* from posterior to anterior. The *left column* shows the positioning of the slices in each series (superior view). The *middle column* shows an example slice (indicated by the *orange arrow*). The *yellow arrowheads* highlight the largest differences of the anatomy relative to the straight coronal slice, which therefore act as the driving force after the initialisation. The *right column* shows the evolution of the mean registration error through the 8 substages. The MRI voxel size was 0.85 mm.

a uniform distribution ($c_1, c_2, c_3, c_4 \in [0, 0.1]$), but thereafter kept constant for all slices. Displacements $u_3(\hat{\mathbf{x}})$ were finally scaled to make the curvature realistic, such that the largest value in the 50×50 pixel central region of the slice was 3 mm.

After sampling, the flat 2-D slices were re-initialised randomly within 15 mm from their ground-truth centre. This setting is supposed to mimic the manual or semi-automatic initialisation of brain slice photographs that represent the cross-sectional anatomy of the brain on a (potentially) curvilinear cut surface. For the registration, an extended stage-3 protocol was used, consisting of 8 substages, with substages 3-5 repeated after the affine re-estimation. The registration error (MRE) was measured for each slice after each substage as the mean distance between the pixels of the registered slice and their ground-truth position in the sampling frame. Background pixels were discarded from the measurement as they were masked out during the registration. The results for the different sampling schemes are summarised in Table 5.2 and plotted in Figure 5.5.

	Slice series			
Stage	Straight	Straight curvilinear	Oblique	Oblique curvilinear
0	5.366	4.935	7.397	7.805
1	0.058	1.570	0.012	1.611
2	0.013	1.558	0.004	1.604
3	0.013	1.641	0.004	1.667
4	0.015	0.125	0.008	0.126
5	0.013	2.093	0.004	2.146
6	0.013	2.220	0.004	2.238
7	0.010	0.124	0.009	0.133
8	0.010	2.153	0.004	2.122

Table 5.2: Comparison of the stage-3 registration accuracy on four series of simulated slices. The values of the table represent the mean registration error (in mm) averaged for all 10 slices after each registration substage. (Slice-specific MREs can be seen in Figure 5.5. These were averaged here to identify trends in the MRE, that are independent from the slice geometry.) The MRI voxel size was 0.85 mm. The highlight shows the results from the first and the repeated free-form registration substage for comparison.

In the *straight* series, the initialisation error was uniformly distributed between 2.5 mm and 12 mm with an average MRE of 5.36 mm for the whole series. Despite the large variation, all 10 slices could be registered equally well by the rigid substage

(MRE = 0.058 mm), which was further improved to MRE = 0.014 mm by the affine substage. The registration error stayed roughly constant through the first iteration of the non-linear optimisations until the second free-form deformation (substage 7) achieved further improvement that was conserved by the second affine re-estimation step (MRE = 0.010 mm). These results show that the rigid optimisation converges to the desired optimum, regardless of whether the initialisation is close or further away. It also proves that the non-linear optimisation can help find a better linear alignment with the affine re-estimation step, although in this case the changes in the MRE were almost negligible after the rigid optimisation.

In the *oblique* series, the initialisation error was higher (MRE = 7.39 mm), indicative of the higher difficulty of the registration. As in the previous case, the rigid substage successfully registered all 10 slices (MRE = 0.013 mm), and the affine substage made a further improvement (MRE = 0.004 mm), that was sustained through the rest of the registration stages.

The registration of the *straight curvilinear* and *oblique curvilinear* slices showed a similar trend, which was different from that of the planar sections. On average, the initialisation error was larger for the oblique series (MRE: 4.94 mm vs. 7.81 mm), but this difference completely vanished after the rigid substage (MRE: 1.57 mm vs. 1.61 mm), indicating that the rotation components were accurately estimated for the oblique slices despite their curvature. As expected, the MREs after the rigid alignment were comparable in size to the orthogonal displacements (<3 mm), and neither the affine transformation (MRE: 1.56 mm vs. 1.60 mm), nor the in-plane deformations (MRE: 1.64 mm vs. 1.66 mm) could make any improvement. In fact, the 3rd substage slightly increased the registration error by falsely attributing slice curvature related misalignments to in-plane deformations. These were reverted and successfully converted to orthogonal displacements in the 4th substage, as indicated by the MREs of 0.125 mm and 0.126 mm for the straight curvilinear and oblique curvilinear series, respectively. This is an important result demonstrating that substage 4 can converge on physically realistic curvatures of the brain slices from

an initial planar estimate, and the curvatures can be estimated with subvoxel ($<$0.85 mm) precision.

At the affine re-estimation stage, the MREs returned to the previous level as expected for 8 of the 10 slices. The reason why the errors remained large in the other two cases is because for these 2 slices the brain masks were deliberately ignored in the re-estimation step to demonstrate the suboptimal alignment when the best affine estimate is computed for the whole slice as opposed to the brain area only. Substage 7 (free-form deformation) successfully registered all slices again with no significant improvement ($\leq 10^{-3}$ mm) over substage 4, followed by the second affine re-estimation step (substage 8), which yielded nearly identical results as the first one for all slices. The repeating trends after substage 5 provide further underpinning that the deformations in the free-form substage(s) are not artefactual, as they converge on the same solution from multiple initialisations.

The main advantage of using simulated slices was that the ground truth coordinates of the registered slices were known, allowing quantification of the registration error based on all brain pixels, in a way that only depended on the accuracy of the registration. On the other hand, the most important limitation was that every slice was an exact match to the MRI data. Consequently, inter-modality contrast differences could not be modelled. This necessitates further experiments to elucidate their effect on the registration performance. Such an experiment is covered in the next section.

5.6 Evaluating the MIND metric versus mutual information

To address the limitation of the previous experiment, the accuracy of registering the *oblique curvilinear* series was evaluated in the presence of added Gaussian noise at each pixel, using the same 8-stage registration protocol. The standard deviation of the noise was adjusted to achieve a consistent SNR of 10 for each slice inside the brain mask. This SNR value is most likely lower than what would be realistic for an actual study. It was picked empirically such that the gray-white boundary is still visible on

most parts of the slices, but concealed enough that an accurate segmentation would be hard or impossible to produce by hand. Consequently, the SNR value corresponds to a level of difficulty that is close to the human limit for this alignment problem.

By default, the stage-3 protocol uses the MIND cost function, which is sensitive to local intensity variations, therefore it may be influenced by the noise. This warrants a comparison with the more popular NMI metric, that may be more lenient to signal variations due to intensity binning. We therefore also registered the noise-free and the noisy variants of the *oblique curvilinear* series using NMI with 32 bins. Table 5.3 and Figure 5.6 summarise the measurements of the MRE in each of the four cases, for each slice, after each substage of the registration.

	Experimental case			
Stage	MIND	MIND SNR=10	NMI	NMI SNR=10
0	7.805	—	—	—
1	1.611	1.725	2.342	3.663
2	1.604	1.767	2.303	3.611
3	1.667	1.795	2.336	3.606
4	0.126	0.575	0.375	2.888
5	2.146	3.010	1.757	3.650
6	2.238	3.109	1.748	3.611
7	0.133	0.513	0.219	2.775
8	2.122	3.039	1.733	3.639

Table 5.3: Comparison of the stage-3 registration accuracy using MIND and NMI on simulated slices with and without Gaussian noise. The values of the table represent the mean registration error (in mm) averaged across all 10 slices after each substage of the registration. The MRI voxel size was 0.85 mm. The two free-form deformation stages are highlighted for comparison.

With the MIND cost, we observed little difference between the average MREs of the *noise-free* and *noisy* series after the rigid registration substage (1.61 mm vs. 1.73 mm, respectively). This shows that the estimation of the slice position and orientation is robust against much weaker contrast correlations than initially tested. However, a marked decline of the registration accuracy was evident after the 4th substage (avg. MRE: 0.126 mm vs. 0.576 mm). While this clearly shows the adverse effect of diminished contrast, the second iteration (7th substage) could

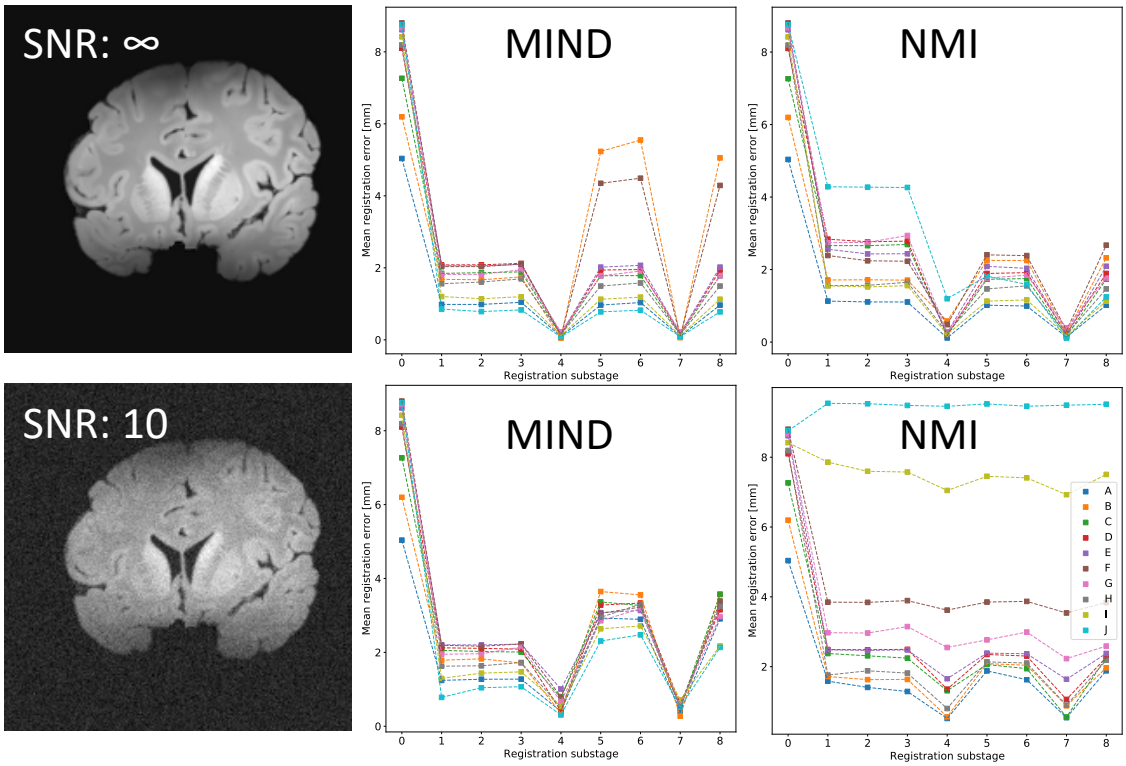

Figure 5.6: Slice-to-volume registration experiment on simulated slices with (x:10°, y:−10°) rotations and additional quadratic through-plane curvature. Registrations are performed with both MIND and NMI used as the similarity metric. The registration accuracy is reported after each substage for both metrics. The experiment is repeated on slices with additional Gaussian noise ($SNR = 10$). MIND outperforms NMI in terms of both robustness and accuracy.

slightly improve this result (avg. MRE = 0.513 mm). Eventually all 10 slices could be registered with sub-voxel precision, which is an important performance indicator considering that the average initialisation error was 7.8 mm). On the noisy slices, both affine re-estimation steps yielded less accurate results than the 2nd substage (avg. MRE: 3.01 mm and 3.04 mm vs. 1.77 mm). This is explained by erroneous brain masking, as the masks were auto-generated from the slice image by thresholding and included noise spikes from the background. One may notice an additional discrepancy between the noisy and the noise-free sets: the MREs of the 2 simulated slices that were previously used to demonstrate the importance of masking, did not produce the same large MREs in the noisy set. This is because masking was consistently applied for all simulated slices in the noisy set.

With the NMI cost, the MREs showed a familiar trend for the *noise-free* slices,

that could be registered with sub-voxel precision by the second iteration of the optimisation. The average MRE at substage 7 was 0.219 mm, slightly higher than the same figure for the MIND cost (0.134 mm). An important qualitative difference between the MIND and NMI based registrations is that the rigid substage was less accurate with NMI (avg. MRE: 1.726 mm vs. 3.663 mm). This was improved after the affine re-estimation, but the average MRE at substage 5 remained slightly higher than for MIND. Notably, the registration with NMI was less robust, as the MRE of one slice failed to get under the voxel resolution on the first iteration.

With NMI, only 3 out of 10 *noisy* slices could be registered with sub-voxel precision. In most cases, the curvature of the slice could not be estimated after the rigid initialisation. In one case, even the rigid substage converged to a more distant optimum. The experiment with SNR = 10 therefore provides strong evidence that NMI cannot differentiate between planar and slightly curvilinear cross sections of the brain as reliably as MIND can. It is therefore recommended to use the MIND similarity metric whenever such deformations of the slices are anticipated.

5.7 Control point optimisation

Non-linear deformations of the 2-D image are optimised at a small set of control points (CPs) distributed across the 2-D image domain. To test how the choice of CPs affects the accuracy and the performance of the registration, we carried out additional experiments on simulated slices. In these experiments we tested (1) different CP placement schemes as well as (2) changing the number of CPs, while registering the noise-free *oblique curvilinear* slice series with the extended stage-3 protocol (8 substages) using the MIND similarity metric. The MREs were measured for each slice after each registration substage, and compared between the test cases.

5.7.1 Control point placement schemes

The following three schemes were tested:

1. **Halton**. Elements of the Halton sequence H_k are generated by the hierarchical k-way division of the $[0, 1]$ interval. [98] The first few elements of the sequence for $k = 1, 2$ are:

$$H_2 = \left\{\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \dots\right\}$$
$$H_3 = \left\{\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \frac{4}{9}, \frac{5}{9}, \dots\right\}$$

A bounding box was generated around the brain area by Otsu-thresholding with +5% margin on each side to contain the CPs. The horizontal and vertical coordinates of the points were then taken from consecutive elements of the H_2 and H_3 Halton sequences, respectively, and scaled by the dimensions of the bounding box.

The Halton method is deterministic and achieves an even coverage of the slice for any number of points, although some of these may lie outside the brain area. Most outliers experience zero registration force from their immediate surroundings, because the background is masked out. Nevertheless, outliers have the tendency to accrue large displacements during the optimisation, because they have a small but persistent remote effect on the nearest brain area (since node weights are calculated globally in Equation 5.3). The magnitude of this effect depends on the properties of the radial basis function, which is Gaussian by default (Equation 5.2). Care must be taken not to interpret these displacement as physical deformations of the slice. A particular advantage of Halton points versus rectangular grids is that the former produces less outliers, because the points are rarely on the edge of the ROI.

Another advantage of the Halton method versus regular grids is that it works for any number of CPs, providing a smoother trade-off between precision and computation time. There is a motivation to use more CPs to capture finer variations of the deformation field, as the support of each CP becomes smaller (it is proportional to the average pairwise distance of the points).

However, every additional CP implies the optimisation of 3 new displacement parameters, rapidly increasing the computational demand.

2. ***k*-means**. The pixel coordinates of the binary brain mask were demeaned, variance-normalised, and classified with uniform weights into a pre-specified number of clusters using the k-means algorithm. Deterministic behaviour of the classification was enforced by fixing the random seed. The resultant cluster means were designated as CPs. Similarly to the Halton method, the clustering can be used to define any number of CPs that are evenly distributed (i.e. devoid of local high- or low-density clusters). Furthermore, the k-means method also guarantees that none of them is outside the brain area. Eliminating outliers means that all CPs predominantly respond to the local deformations of the slice. Compared to the Halton method, the CPs are also denser in the brain area, which means greater capacity to estimate deformations at a finer spatial scale at the same computational cost.

3. **Cost adaptive.** This method also uses k-means clustering to generate CPs. Extra steps are employed to concentrate the points in the most misaligned regions. This strategy may lead to an uneven distribution of CPs, but provides greater flexibility to estimate local deformations of the slice where the demand is the greatest, while using the same number of points. The CPs are generated successively, usually in 4 batches. Let k_0 and k_1 denote the counts of previously defined points and new points, respectively. First the brain area is parcellated into $k_0 + k_1$ regions using k-means clustering, and the image dissimilarity term (cf. Equation 3.7) is mapped over the same area. The dissimilarity map is smoothed within the brain mask with a Gaussian kernel (FWHM $= \frac{2A_{\text{brain}}}{k_0 \pi}$), and the centres of the regions with the k_1 largest mean dissimilarity value are taken as the new CPs, which are added to the existing pool. The CPs are generated in parallel with the optimisation of the non-linear transformation. The first iteration is identical to the k-means method, because $k_0 = 0$. This creates and evenly distributed set of CPs. From the second iteration onwards,

points will start to accumulate in the most misaligned region. On each of these iterations, a new field (`TxRbfDisplacementField`) is created, its parameters are fitted to the previously optimised field (cf. Equation 5.3), allowing the optimisation to proceed from the previous state with additional DOFs where needed. If an optimisation step resolves a misalignment in one region, the CPs in the next iteration will start to concentrate in another region.

A comparison of the three schemes using 16 CPs is shown in Figure 5.7, and the exact MREs are summarised in Table 5.4. The registrations proceeded identically through the 8 substages for all three schemes, in keeping with the previous descriptions of the registration of the *oblique curvilinear slice series*. All results were visually accurate in all three cases at the end of the 4th substage, indicating the resilience of the stage-3 protocol to large variations in CP pattern. By strict comparison of the average MREs, the *Halton* method was found to be superior to the other two, with the *Cost adaptive* method performing slightly better than the native k-means.

	CP placement scheme		
Stage	k-means	Halton	Cost adaptive
0	7.805	—	—
1	1.611	1.597	1.597
2	1.604	1.601	1.601
3	1.667	1.651	1.661
4	0.126	0.057	0.104
5	2.146	2.098	2.019
6	2.238	2.184	2.107
7	0.133	0.058	0.111
8	2.122	2.115	1.988

Table 5.4: Comparison of stage-3 registration accuracy using different control point placement schemes. The values of the table represent the mean registration error (in mm) averaged across all 10 slices after each substage (0-8). The results of the two free-form deformation substages are highlighted for comparison.

Although it cannot be fully dismissed, it is unlikely that the superior performance of the Halton method is linked to some special mathematical property of the Halton sequence, as the very different pattern of the other two methods did not give

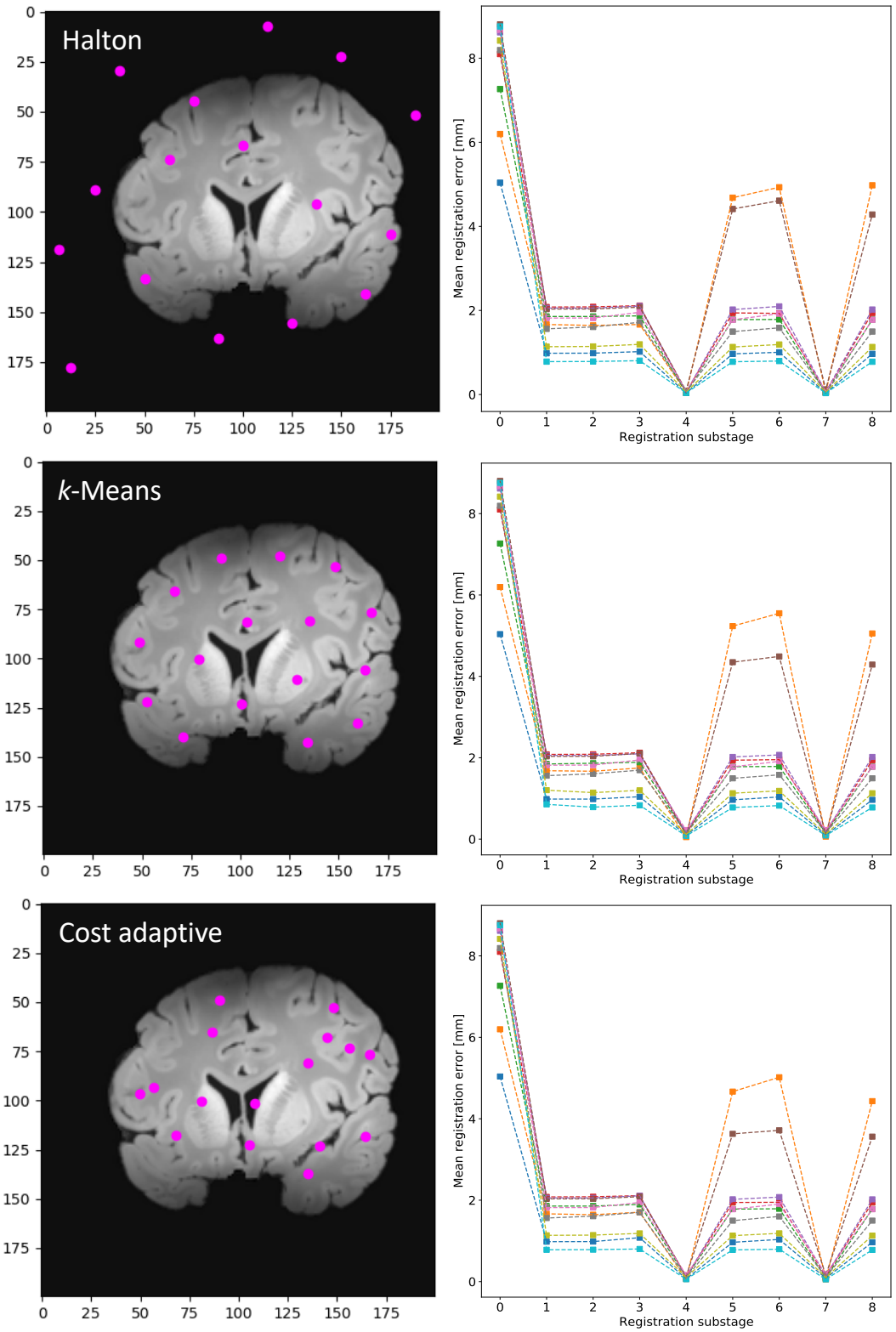

Figure 5.7: Evaluation of the registration accuracy on oblique simulated slices with additional through-plane curvature using different control point placement schemes.

rise to a comparable difference in the MREs. Visualising the distribution of the registration error on an example slice (Figure 5.8) points towards a more specific explanation. (1) The points generated by clustering did not extend beyond the brain area, and the curvature of the slice was therefore incorrectly extrapolated beyond the outermost CPs, which inflated the errors on the slice edge. (2) Cluster points were denser than Halton points, which implies they also had a smaller support. If the supports are too small to overlap, the registration may become inaccurate midway between the points, but also at the CPs. As the support radius tends to zero, the local curvature of the slice vanishes, and the sensitivity of the CP shifts from the large-scale anatomical differences that should drive the orthogonal deformations to the more local intensity variations, which can therefore give rise to fluctuations around an arbitrary z-displacement. These effects are seen in the stronger ripples of the registered slice with the k-means CPs, although these tend to be smaller than the voxel size (0.85 mm).

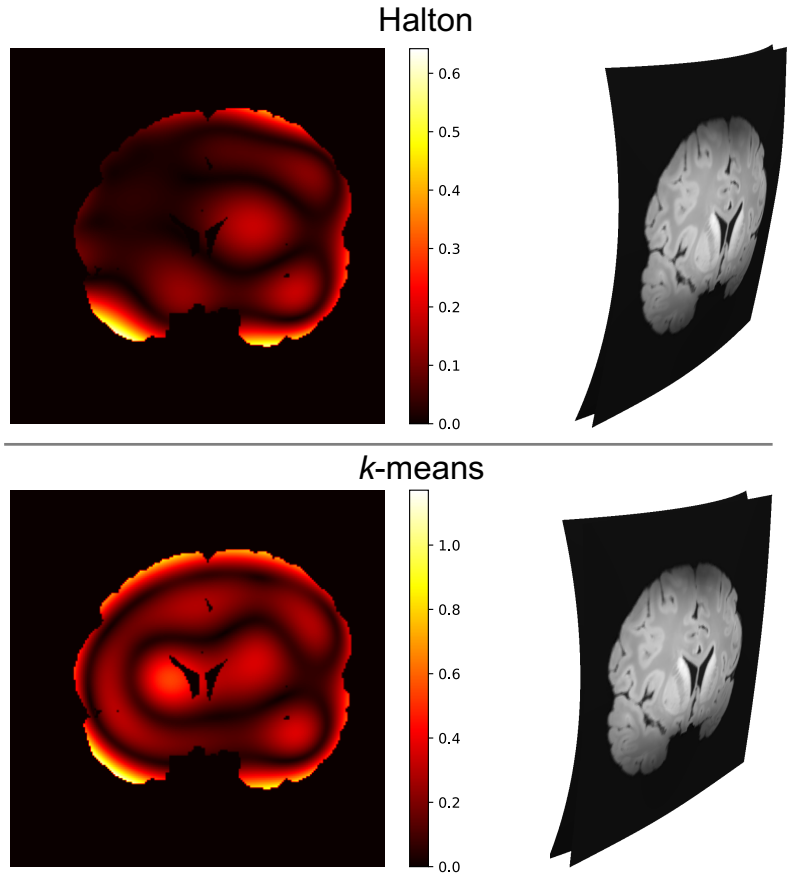

Figure 5.8: Spatial distribution of registration errors (in mm) on an example oblique curvilinear slice after the 4th registration substage with *Halton* and k-means CPs. The 3-D reconstructions show the alignment of the registered slice relative to the ground truth. In both cases, the curvature is only estimated correctly under the coverage of the CPs. The difference of the curvatures is slightly larger with the k-means points.

5.7.2 Number of control points

The extended stage-3 pipeline (8 substages, MIND metric) was run on the simulated *oblique curvilinear slices* using 4, 8, 16, 32, 64, and 96 CPs generated by the k-means method. The k-means method was chosen for its property to create the most uniform spatial distribution for any number of points, and because it was shown in the previous experiment to provide an upper estimate of the registration error for a given number of points among the three methods. For a fair comparison, the rigid and affine substages of the pipeline were executed only once, and used as a common initialisation for all test cases. Figure 5.9 shows the registration curves, and the corresponding average MREs are reported in Table 5.5.

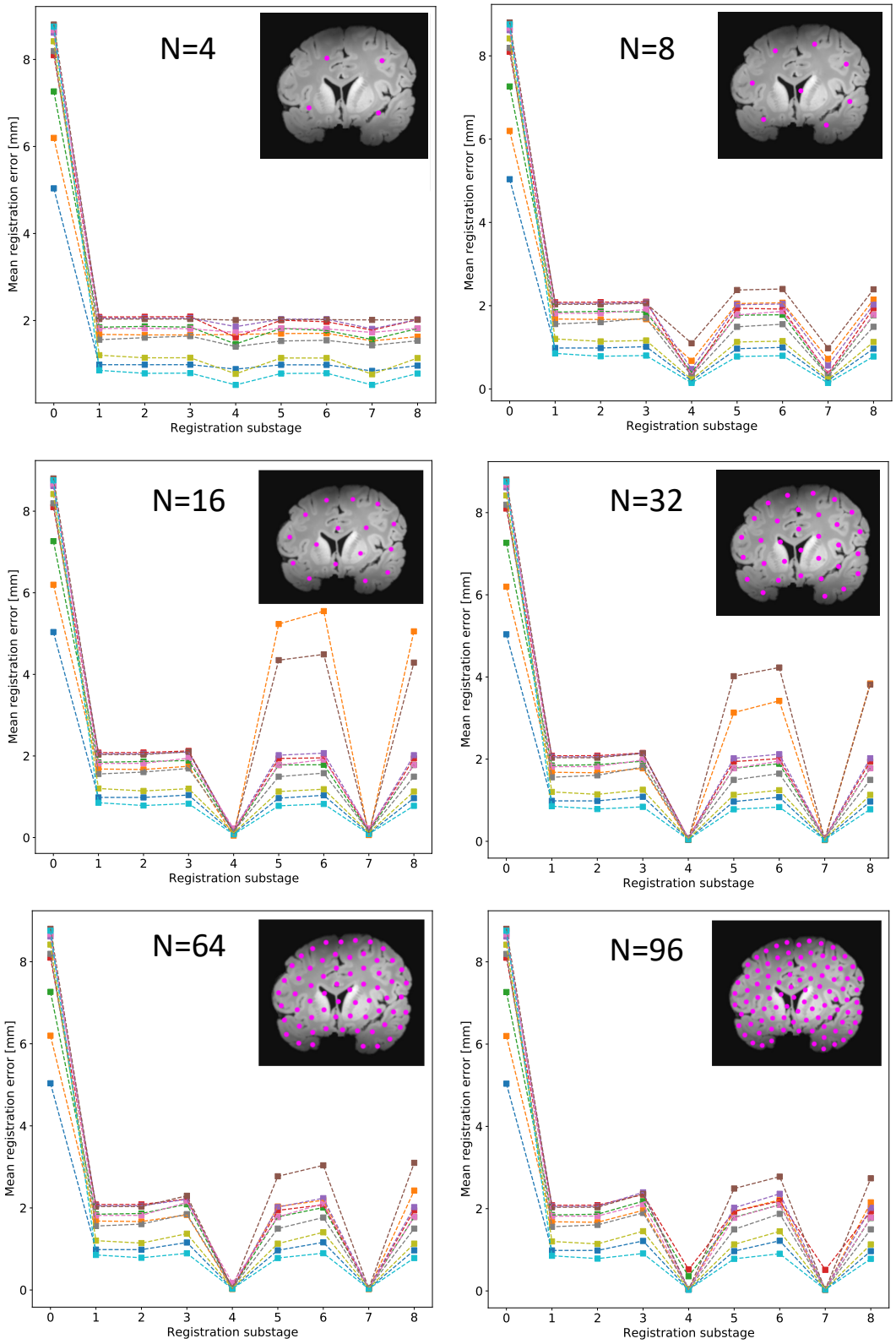

Figure 5.9: Evaluation of the registration accuracy on oblique simulated slices with additional through-plane curvature using different number of control points ($4 \leq N \leq 96$).

	Number of control points						
Stage	4	8	16	32	48	64	96
0	—	—	7.805	—	—	—	—
1	—	—	1.611	—	—	—	—
2	—	—	1.604	—	—	—	—
3	1.607	1.634	1.667	1.713	1.765	1.805	1.881
4	1.392	0.422	0.126	0.058	0.046	0.052	0.118
5	1.582	1.630	2.146	1.904	1.807	1.670	1.632
6	1.575	1.658	2.238	2.041	1.983	1.888	1.916
7	1.397	0.421	0.133	0.060	0.045	0.039	0.084
8	1.574	1.642	2.122	1.954	1.845	1.741	1.677

Table 5.5: Comparison of stage-3 registration accuracy using different numbers of control points. The values of the table represent the mean registration error (in mm) averaged across all 10 slices after each substage (0-8). The results of the two free-form deformation substages are highlighted for comparison.

The difference between the 4th and 7th substage results were mostly negligible, indicating that the affine initialisation of the slices were accurate, and the non-linear stages converged towards the same optima, close to the ground truth. At the end of the 7th substage, the registration errors monotonically decreased in the 4–32 point range, then plateaued with a moderate increase at the final 96 point mark.

The results were visually accurate for all slices with 16 or more points. Fewer CPs (4, 8) systematically underestimated the orthogonal deformations of the slices, although a slight improvement could be observed with both 4 and 8 points relative to the best affine result. At 96 points, the parameter space became excessively high-dimensional, and the optimisations most likely converged on local optima for both the in-plane and free-form deformations, causing a decline of the registration accuracy. The monotonically increasing MREs at the end of the 3rd substage indicate the greater prevalence of spurious in-plane deformations with more CPs, which provided a worse initialisation for the later substages, further contributing to this effect.

Based on the above, the MREs suggest a wide optimum range (16-96), in which the registration accuracy is relatively insensitive to the number of CPs used, leaving the choice of CPs up to considerations of runtime. Table 5.6 lists the runtimes of the

Stage	Number of control points						
	4	8	16	32	48	64	96
1	—	—	438	—	—	—	—
2	—	—	28	—	—	—	—
3	13	49	150	858	1103	1803	4298
4	29	101	243	1316	1648	2814	5308
5	0.08	0.10	0.09	0.70	0.40	0.50	0.80
6	13	49	144	890	1135	1946	3382
7	29	107	244	1344	1537	2236	4160
8	0.06	0.10	0.10	0.6	0.33	0.39	0.65
total [min]	9.2	12.9	20.8	81.3	98.2	154.5	293.6

Table 5.6: Average runtimes of the substages in seconds as a function of control points.

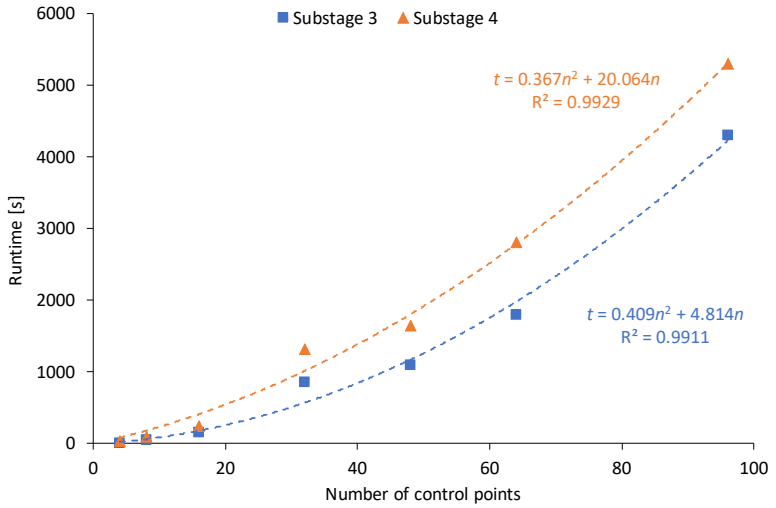

Figure 5.10: The runtime of the non-linear substages 3 and 4 exhibits a quadratic dependence on the number of CPs.

substages for different numbers of CPs in seconds. The values represent averages of 10 slices at 0.85 mm MRI resolution, as measured on a workstation with 2 Intel Xeon X5670 CPUs (12 hyperthreading cores at 2.93 GHz). The runtimes of the rigid and affine substages represent multi-core performance using all 24 logical compute nodes, whereas the non-linear registrations were allocated to single compute nodes and carried out simultaneously. The runtimes of the in-plane and the free-form deformation substages scaled quadratically with the number of CPs. (Figure 5.10). For the registration of the simulated oblique curvilinear slices 16 CPs seemed to provide the best compromise between accuracy and runtime.

5.8 Registration of a real brain slice photograph

To further demonstrate the accuracy of the stage-3 pipeline and the need for orthogonal deformations, we registered an actual brain slice photograph (not a simulated slice) to the corresponding structural MRI volume (Figure 5.11). These images were taken from another subject, not the one whose data was used for the simulations. The photograph showed the posterior surface of a coronal section through the *calcar avis* and the occipital horn of the lateral ventricles. The free-hand cut left the hemispheres detached, and based on the asymmetric appearance of the ventricles, it was slightly oblique. After applying the pre-processing steps as discussed in Section 5.2, the extended stage-3 pipeline was run with default configurations, as outlined in Section 5.4. Namely, the rigid substage tested the combinations of 11 positions in a 20-mm thick search region with 25 orientations within $\pm 20^\circ$ from the estimated physical slicing orientation. For the non-linear stages 16 points were generated by the k-means scheme, and these were optimised in 4 batches at 1.7 mm resolution, and finally once at 0.85 mm. The accuracy of the registration was assessed qualitatively: the gray-white matter boundary was manually annotated on the brain slice photograph using ImageJ, and the contour was overlaid on top of the resampled MRI-slices after each substage (2: affine, 3: in-plane, 4: free-form).

The linear substages estimated the obliqueness of the cut surface accurately enough to reproduce the gross shape of the hemispherical cross sections with the asymmetric appearance of the lateral ventricles. However, a closer inspection of the reconstructed MRI slice with the contours reveals several regions where the registration was less accurate (*yellow arrowheads* in Figure 5.11). In most of these regions, tissues are not simply misaligned, but appear anatomically different, which is a hallmark of orthogonal deformations. Accordingly, the misalignments in these regions persisted through the in-plane registration substage, and could only be eliminated in the free-form deformation substage, producing an almost perfect alignment of the gray-white contours. As shown in Figure 5.11, the through-plane deformations were larger than the voxel size for the majority of the slice area, and

even reached as far as 6.5 mm away from the best affine plane. Meanwhile the Jacobians showed a conservative ±10% dilation/shrinkage of the slice throughout the entire 2D/3D transformation, indicating the non-linear deformations were predominantly related to slice curvature. The spatial distribution of through-plane deformations appears physically realistic, resembling the bulging of the core of the hemispheres (an elastic medium) that is compressed by an external force. While assessing the plausibility of the estimated shape of the cut surface, one must also take into account potential variations of the knife path and the inhomogeneous structure of the brain (especially around the lateral ventricles).

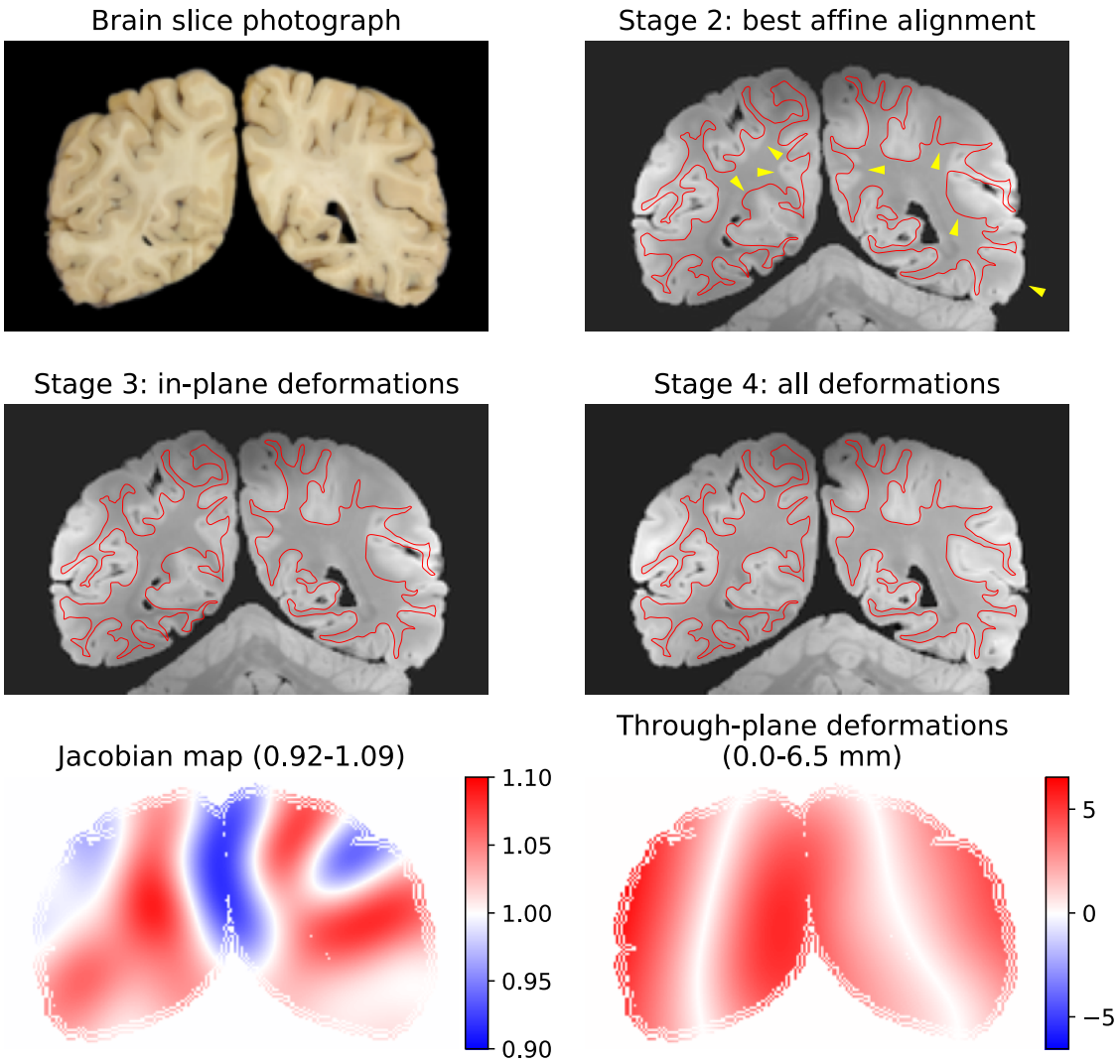

Figure 5.11: Slice-to-volume registration result of an actual coronal brain slice photograph. The *red contour* was hand-drawn on the brain slice photograph and is provided to assess the registration accuracy. The *yellow arrowheads* indicate regions in the resampled MRI slice where the anatomy is most different from that in the tissue block photograph. Through-plane deformations prove to be essential for this slice. The corresponding Jacobian and through-plane deformation maps are shown in the *bottom row*. See main text for further details.

5.9 Comparison with manual landmarks

In the previous section we used gray-white matter contours to demonstrate that the automated stage-3 pipeline can accurately register a real brain slice photograph to volumetric MRI data. In this section we expand on this demonstration by attempting to measure the accuracy of the registration using manual landmarks.

For this purpose, a different coronal section was randomly selected from another

subject, and registered with the same method as described in the previous section. Before the registration, we identified 20 salient points in the brain slice photograph, and manually located each one of them in MRI space by painstakingly scrolling back and forth between the image slices and looking for maximum anatomical similarity. After the registration, the pairwise distances of the transformed slice points and the MRI points were observed to follow a chi-distribution with a median of 0.93 mm and an interquartile range of 0.37–1.31 mm.

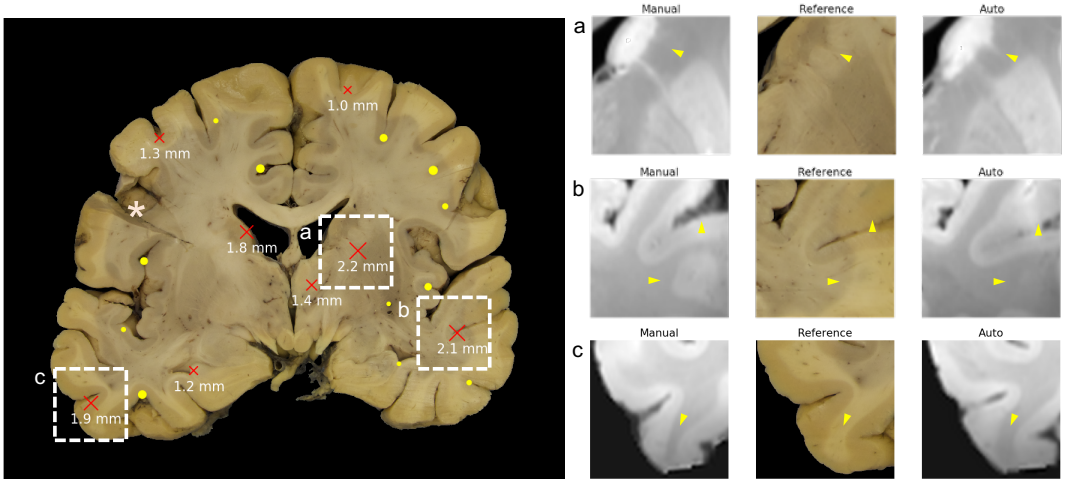

Figure 5.12: Evaluation of slice-to-volume registration errors on a real brain slice photograph, using anatomical landmarks. *Yellow dots* indicate landmarks within 1 mm from the registered slice, *red crosses* indicate landmarks further away. (**a**, **b**, **c**) compares landmark-based and automated registration in the regions with the largest error. Automated slice-to-volume registration is clearly superior in these regions, indicating relatively large errors in the annotation.

Previously (in Section 5.1) we argued that finding the position of a 2-D slice in volumetric data by point pairs is prone to errors, especially where the anatomy is fairly consistent across several consecutive MRI slices. To investigate the annotation bias in the measured distances, we visualised the alignment of the resampled MRI slice with the photograph around each landmark, and further compared these with an MRI slice that was resampled using the landmarks. For the latter visualisation, an alternative stage-3 transformation chain was constructed with rigid parameters that initialised the sampling frame as the best planar fit to the landmarks. The planar projection of the landmarks were then designated as CPs for the non-linear

transformation, and the orthogonal displacements of the frame were calculated by fitting Gaussian radial basis functions to the position of the landmarks (as in Equation 5.3). The anatomical features of the resampled MRI slices were qualitatively compared with those of the photograph. Where the position of the landmark is accurate, the expectation is that it is the landmark-based MRI snippet that resembles the photograph more.

Figure 5.12 shows the distance of the landmarks from the automatically registered slice. Comparisons between the automated and the landmark-based alignment is shown for the three regions where the result of the stage-3 registration was the furthest away from the landmarks. Surprisingly, in all three cases presented, and altogether in 6 of the 8 regions where the apparent registration error was larger than 1 mm, the automated alignment was consistently more accurate than the landmark-based alignment. Counterexamples, where the accuracy of the automated method was slightly inferior to that of the manual landmarks were found in only two cases: one in the proximity of the damaged area (*pink asterisk*), and the other where the side surface of the brain slice was visible and locally biased the registration towards larger out-of-plane distortions. The accuracy of slice-to-volume registration in these regions should be improved by customising the slice mask to exclude damaged areas and visible side surfaces.

These findings are important, because they show that manually matching MRI slices to 2-D images by visual comparison can be rather inaccurate despite careful practice. Nevertheless it is seen as common practice where suitable software/hardware solutions for accurate MRI-histology registration are not readily available. When several slices are considered, the automated method is not only faster, but also more accurate, and the result can be traced back to find exact coordinates in both image spaces.

6

Histology-to-MRI Registration Experiments with TIRL

Contents

6.1 Registration by two photo intermediates 120
6.2 Registration by one photo intermediate 150

In the previous chapters, we described the development and testing of three separate registration routines within the novel TIRL framework: *stage 1* (histology to tissue block photograph), *stage 2* (tissue block to brain slice), and *stage 3* (brain slice to volumetric MR data). In this chapter we will demonstrate how these routines are combined to provide end-to-end registration between microscopy and MR images in various settings.

In the previous chapters we argued that tissue photographs are useful intermediate elements of the pipeline. Specifically, (1) the *tissue block photograph* serves as a quasi-rigid reference, that allows more accurate estimation of the deformations introduced by histological tissue processing than MRI alone, and (2) the unique cross-sectional anatomy of the *brain slice photograph* allows high-precision 3-D localisation, especially when the slicing surface is tilted or curved. However, photographic intermediaries may not always be available even where MRI and

microscopy images are jointly acquired. Therefore this and the next chapter are partly dedicated to explore registration approaches that use different numbers of photographic intermediates to solve real-life problems: two intermediates in Section 6.1, a single intermediate in Section 6.2, and finally some preliminary results with no intermediates in Section 7.2.2. Each case is presented with a separate description of the data and the methods. Results are provided to evaluate the versatility and resilience of the TIRL framework in each of these settings.

6.1 Registration by two photo intermediates

6.1.1 Dataset

This demonstration is carried out on the post-mortem human ALS cohort, that inspired the development of the TIRL framework and was used throughout the testing of its registration routines. The collection of the MRI, photographic and histological data for this cohort was described in Section 3.1 and also in [21]. Here we only draw attention to a few details that are important for the purpose of this demonstration.

The cohort consisted of the post-mortem brains of 13 ALS patients and 3 controls. Each of the brains underwent multi-modal post-mortem MRI scanning with subsequent dissection and histological processing. One individual's dataset was discarded for an incomplete set of dissection photographs, and another one for severe susceptibility artefacts in the structural MR image. In the remaining 14 brains, only a specific subset of the tissue blocks were considered for registration, where the following conditions were jointly satisfied:

- Photographs of both the tissue block and the corresponding brain slice photograph were available to facilitate the registration between histology and MRI. This includes the majority of the blocks, which were sampled by excision from the coronal sections, and excludes the motor cortical specimens that were sampled directly from the convex surface of the hemispheres before the coronal sectioning took place.

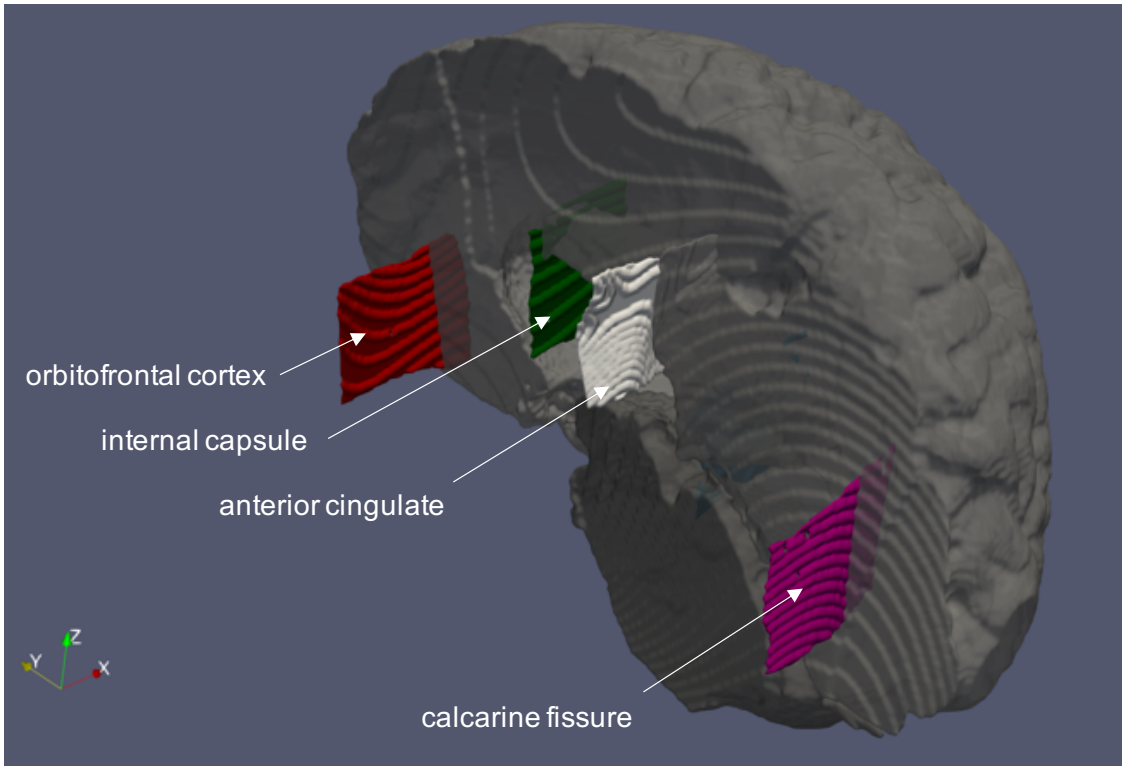

Figure 6.1: Three-dimensional model of the post-mortem brain showing a subset of the histological sampling sites, as they were reconstructed from the registration results. Each site is represented by the curvilinear domain of the corresponding histology image, and is therefore larger than the actual tissue section. The left-hand side of the brain was removed for better visualisation, some of the sampling sites are also not shown for greater clarity. Note that certain curves are accented because the surfaces were reconstructed from voxel-wise labels in MRI space (voxel size: 0.85 mm).

- Immunohistochemical staining for myelin proteolipid protein (PLP) had taken place to produce histology images with suitable gray-white matter contrast.

As a result of this selection process, 6 suitable anatomical regions were identified that were consistent across all 14 brains: (1, 2) primary visual cortex (left and right), (3) body of the corpus callosum, (4) anterior cingulate cortex, (5, 6) hippocampus (left and right).

For one of the 14 brains, histology was carried out on 4 additional blocks, to allow testing the registration pipeline in a more diverse set of anatomical regions: (1) orbitofrontal cortex (left), (2) Broca's homologue area (right), (3) thalamus (bilateral), and (4) internal capsule (right). Figure 6.1 illustrates a subset of these sampling locations in 3-D.

In preparation for the registration, a table of corresponding images was created, in which each histology image was linked to a specific set of photographs and an MR volume. The photographs were all background-segmented (see Section 3.5), and included both the tissue block photograph, the intact brain slice, and a series of photographs of the cut-out brain slice. Care was taken to identify whether the histology image was sampled from the posterior (95%) or the anterior (5%) surface of the blocks, and equally, whether this was correspondent to the posterior surface of a specific brain slice, or the visually identical but mirrored anterior surface of the next slice. Where the anterior surface was sampled, the pixel order of the histology image and the photographs was flipped along the first image axis to ensure that all images were uniformly represented in "neurological" orientation. The structural MR volumes were resampled to 0.25 mm isotropic resolution for each subject using FLIRT [99], and also linked to the histology images. All subsequent registration steps were carried out on these set of images. Stage-specific parameter configurations were partly different from what was used in the previous chapters, hence these are summarised in the next section.

6.1.2 Methods

Stage 1, 2 and 3 registrations were carried out on a total of 88 histology images and the concomitant photographs and MRI volumes. The resultant transformation chains from each stage were combined to obtain a mapping from histology pixel space to MRI voxel space. To visualise the alignment, the pixel coordinates of the histology image were transformed by the combined chain and the MR data was resampled by trilinear interpolation at the transformed coordinates. The resolution of the histology images was set to 8 µm/px, which corresponds to 16x downsampling from their native resolution. Figure 6.2 shows the registration process on a visual cortical specimen in the left hemisphere. Image frames of identical colour represent identical image domains, where both the logical image shape, and the image resolution are identical, hence the pixels are in 1:1 correspondence. In the following paragraphs we discuss the typical configurations that were used in each stage of the pipeline.

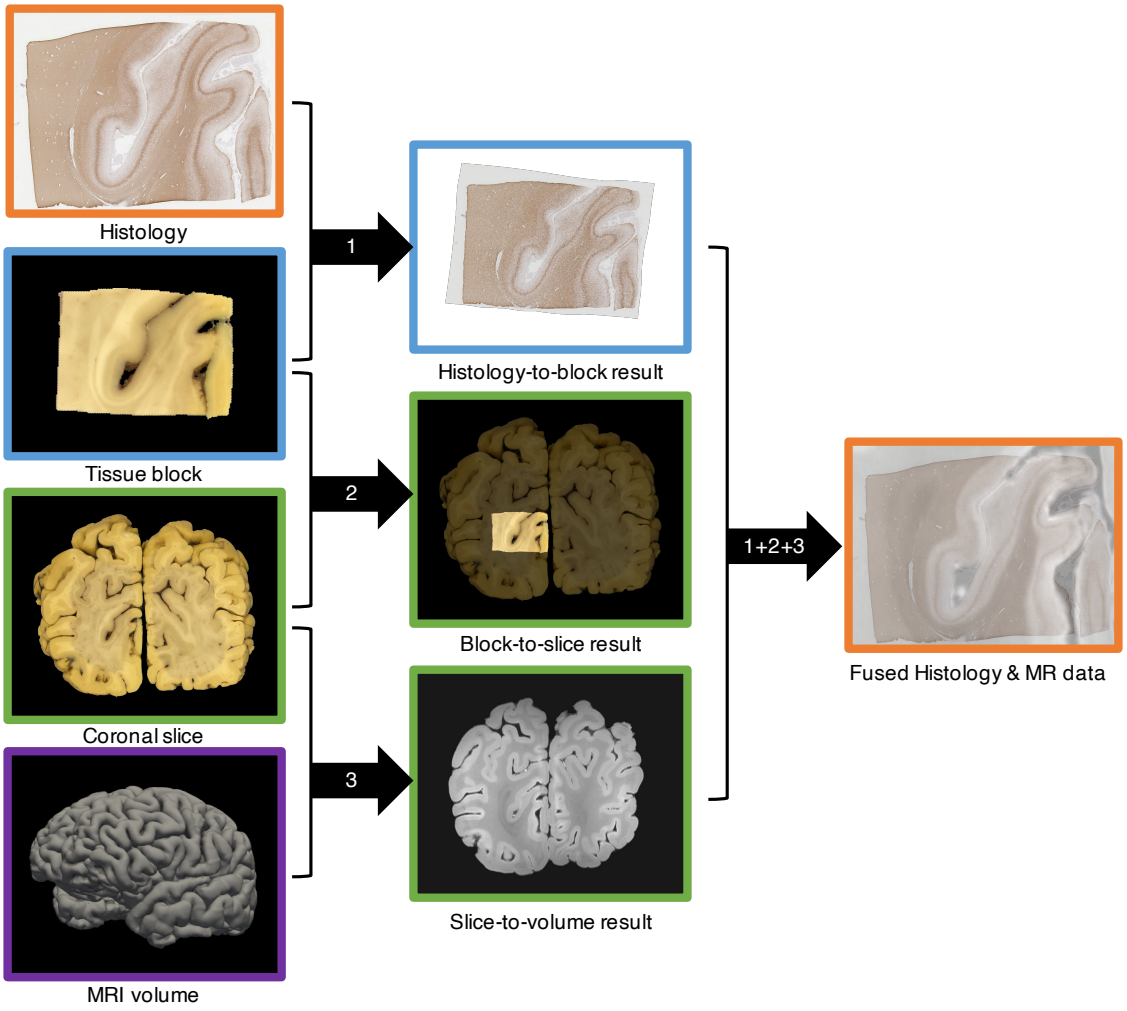

Figure 6.2: Overview of the combined end-to-end histology-MRI registration protocol using two photographic intermediates: the tissue block and the brain slice photographs. Identical frame colours indicate identical image domains with pixelwise correspondence. Stages 1-3 are run in parallel, and the optimised transformation chains from each are combined to transform the histology domain into MRI space, and resample the volumetric MR data onto it. The 2-D MR image is shown here as an overlay on the histolgy image with 50% opacity.

Stage 1 configuration

The PLP-stained histology images were resampled to match the resolution of the tissue block photographs, which was set to a uniform estimate of 50 µm/px. The histology images were converted to grayscale using the YIQ method, and padded by $1/6^{\text{th}}$ of the image size in all directions. No masks were defined for the histology images. The block photographs were also converted to grayscale. A mask was generated automatically by thresholding at 10% of the maximum intensity value (to

exclude the black background), and dilating this segmentation using a 10×10 kernel. After a coarse rotation search with 30-degree increments at $1/8^{\text{th}}$ scale, rigid registration was performed at $1/16^{\text{th}}$, $1/8^{\text{th}}$, and $1/4^{\text{th}}$ scale, followed by affine registration at $1/8^{\text{th}}$ and $1/4^{\text{th}}$ scales, without additional smoothing at any of these resolutions. Bounds on the transformation parameters were ±5 mm for the translation, ±10% for the scale, ±15° for the rotation, and the following values for the elements of the 2-D affine matrix (in row-major order): $\pm \begin{bmatrix} 0.05 & 0.2 & 1.0 & 0.2 & 0.05 & 1.0 \end{bmatrix}$. The linear alignment was fine-tuned using diffusion-regularised deformable registration using 20 iterations at $1/16^{\text{th}}$, $1/8^{\text{th}}$ and $1/4^{\text{th}}$ scales, with a regularisation weight of 0.4. The more conservative regularisation was chosen to minimise spurious deformations, in keeping with our previous experiments that indicated comparable registration accuracy at weights 0.3 and 0.4. All registration steps minimised the MIND cost metric, where the descriptors were computed for a dense 3×3 neighbourhood with Gaussian smoothing radius $\sigma = 1$.

The results of the registrations were carefully observed to identify potential quality issues. We used a purpose-built interactive visualisation tool that displayed the registered images on top of each other. The opacity of the top image was gradually changed, and we looked for "jumps" in the anatomical pattern as one of the images gradually faded into the other.

Stage 2 configuration

Possible insertion sites were automatically inferred for each brain slice by running the TIRL site finder script on each corresponding series of intact and cut-out brain slice photographs, as described in Chapter 4. For each block insertion, the tissue block photograph and the intact brain slice photograph were converted to grayscale using the YIQ method, and their resolutions were uniformly set to 50 µm/px. Cost function masks were created for both images by thresholding at 10% of the maximum intensity value (to exclude black backgrounds). The previously determined insertion sites were fed into the configurations of the stage-2 registration script.

During stage-2 initialisation, the alignment of the tissue block was evaluated by the MIND metric at $1/10^{\text{th}}$ scale. A grid search was performed over the rotation and translation parameters, such that 12 rotations were tested with 30° increments, combined with translations in each direction by $\begin{bmatrix}-10 & 0 & 10\end{bmatrix}$ mm (altogether $12 \times 3 \times 3 = 108$ parameters combinations). The best 3 results at each site were fine-tuned using BOBYQA optimisation, and the parameters associated with the lowest cost were used to initialise the similarity (rigid + scaling), affine, and diffusion-regularised non-linear registration steps.

The subsequent steps of stage-2 registration were carried out at multiple resolution levels: the similarity transform at 1/10 and 1/5, the affine at 1/5, 1/2 and native resolution, and the non-linear at 1/20, 1/10, and 1/5 scales. At each of these, BOBYQA optimisation was used to minimise the MIND metric (dense 3×3 neighbourhood, $\sigma = 1\,\text{mm}$), and the respective parameter bounds were ±15° for the rotation, ±5 mm for the translation, and the following for the affine parameters (in row-major order): $\pm\begin{bmatrix}0.2 & 1.0 & 5.0 & 1.0 & 0.2 & 5.0\end{bmatrix}$. The non-linear registration employed 20 iterations at each resolution level, and a regularisation weight of 0.6 was used without bounds on the displacement vectors. As discussed in Chapter 4, the higher regularisation weight was a reasonable compromise of flexibility, as the anticipated deformation of the tissue block between the two images was small.

Stage 3 configuration

With respect to slice-to-volume registrations, there was a redundancy in the input data, such that multiple blocks (especially left-right pairs) had been extracted from the same brain slice, and these slices should be registered to the respective MR volume only once. As stage-3 registration requires an initial position and orientation estimate for each slice, a helper script and a batch processing script were created to simplify the provision of these estimates for slices that are pertinent to the same subject. For each subject, the 6^{th}coronal slice from the anterior pole was designated as the *reference slice*, as it is the most feature-rich, and the cross sectional shape of the lateral ventricles normally provides a reasonable guide for estimating even some

of the less intuitive rotations. The respective MR volume was loaded in FSLeyes, and the cursor was placed at the approximate centre of the reference slice. The Nudge tool was then used to estimate the rotations about each axis, centred at the cursor location. Based on the cursor location and the rotation angles, the helper script estimated the position of other slices by shifting the centre of the reference slice along its normal vector according to the slice number and the average 10 mm thickness of the slices. While the majority of these estimates were accurate within the expected ±5 mm range, the positions had to be fine-tuned in about 25% of the cases to account for variabilities in slice thickness. For fine-tuning, the helper script resamples the MR data at the estimated position of the slice, displays it next to the photograph, and provides a slider to adjust the distance of the slice from the reference slice. After all adjustments were made, the helper script created a batch configuration file (one per subject), which was used by the batch processing script to create individual stage-3 configuration files and target directories for the registration of the individual slices. Altogether 209 brain slice photographs were registered for 14 subjects, corresponding to an average of 15 slices/subject (the number of slices depended on the size of the brain, and a few slices were registered with both their anterior and posterior surfaces).

In each case, stage-3 registration was performed with 5 substages on the background-segmented intact brain slice photographs and the structural MR volume (TRUFI, 0.25 mm isotropic): (1) similarity transform (with 2-D and 3-D substages), (2) affine transform (3-D), (3) in-plane deformation (2-D), (4) free-form deformation (3-D), and (5) affine re-estimation. The photograph was converted to grayscale with the YIQ method and resampled to match the resolution of the MR volume. Both images were normalised to the maximum intensity value.

Masks were defined automatically for the brain slice photograph by thresholding at 10%. The automatically generated masks were occasionally manually adjusted to exclude visible gyri that were not part of the cut surface using ImageJ. The cerebrum mask generated by BrainSuite was specified as a mask for the MRI volume to exclude the cerebellum from the registration.

All registration steps employed BOBYQA optimisation to minimise the MIND dissimilarity metric (computed for a cross-shaped 3×3 neighbourhood with $\sigma = 1$). The resolution of the photographs were inferred from ruler readings and were supplied through the batch processing script.

Substage 1 (similarity transform) was configured to explore 5 positions in a 10 mm thick rectangular search region, in combination with $1 \times 5 \times 5$ different z-y-x intrinsic rotations in the following range: $\pm \begin{bmatrix} 0° & 10° & 10° \end{bmatrix}$ (in-plane rotations were fine-tuned within the same $\pm 10°$ bounds after each initialisation during the 2-D optimisation). All initialisations (altogether 125) were tested at 1/8 scale, then the best 30% (i.e. 37) were tested again at 1/4, followed by the refinements of the best 11 initialisations at 1/2 and finally 3 initialisations at native resolution. No additional smoothing was employed at any of these scales. During substage 1, the bounds were ± 10 mm on the 2-D translation parameters, ± 3 mm on the 3-D translation parameters, and scaling was restricted to $\pm 5\%$.

In substage 2, the affine registration was carried out at scales 1/4, 1/2, and 1, using ± 1 mm bounds on the translation parameters and ± 0.1 on the rest of the affine parameters.

The deformable registration in substages 3 and 4 were performed using 32 control points with the Halton placement scheme. The Halton scheme was chosen for its superior performance versus the other two schemes. The larger number of CPs were chosen to provide additional flexibility for the transformation to fully utilise the higher resolution of the structural MRI data. The respective *xy* and *xyz* displacements were optimised in increments at 1/2 scale: first for 8, then for 16, 24, and finally for 32 points. The 32-point registration was further refined at native resolution (0.25 mm). Bounds on the displacements were set to ± 5 mm on each axis. No additional regularisation was found to be necessary in either substage.

Combining stages 1, 2, and 3

The primary outputs from each registration stage are the optimised transformation chains that map each of the input images onto a common image domain that is

embedded in physical space. Detailed descriptions of the stage-specific transformation chains were given in the previous chapters, and we refer to Figures 3.4, 4.3, and 5.2 for an overview of each.

To obtain a pixel-to-voxel mapping between corresponding pairs of histology images and MR volumes, the following transformation chains are concatenated from left to right, i.e. transformations are applied to the histology coordinates in the order as they appear in the combined chain from left to right:

1. Stage 1: histology pixels at 8 µm/px to physical space: $S_1 \rightarrow S_4$
2. Stage 1: physical space to tissue block pixels at 50 µm/px: $S_4 \rightarrow D_5$
3. Stage 2: tissue block pixels at 50 µm/px to physical space: $D_0 \rightarrow S_3$
4. Stage 2: physical space to brain slice pixels at 50 µm/px: $S_3 \rightarrow D_4$
5. Stage 3: brain slice pixels at 50 µm/px to physical space: $D_0 \rightarrow S_4$
6. Stage 3: physical space to MRI voxels at 0.25 mm/vx (inverse *sform* matrix): $S_4 \rightarrow D_5$

It should be noted that the combination ignores the internal part of the stage-1 chain, hence the combined mapping is from S_1 space, not D_0. This corresponds to the pixels of the original histology image at 8 µm/px resolution, not the pixels of the padded and downscaled version of the histology image that was actually used in stage 1 to estimate the transformation parameters. This exemplifies the use of internal transformations in TIRL: temporary or stage-specific image operations can be kept separate, and removed from the chain very easily when they become unnecessary. If these were baked into a single affine transformation together with the linear part of the chain (as it would be in the case of a NIfTI image), it would be much less intuitive to undo these when the images are used in a different context, e.g. at a different resolution, or without padding. If the original histology image is loaded at the maximum 0.5 µm/px resolution, TIRL fills the internal stack with a new scaling transformation that maps the large 0.5 µm/px image

to the pixel grid that corresponds to 8 µm/px, and the combined chain remains a valid mapping into MRI space.

Before the combination, each non-linear transformation is dynamically linked to the antecedent part of the stage-specific transformation chains. This means that the domain on which the displacements are defined also move when any of the antecedent transformation parameters are changed. After the combination, the non-linear transformations become statically defined, e.g. the non-linear transformation of stage 2 no longer moves if any of the linear transformations from stage 1 are changed. Should those parameters change, the mapping would become erroneous, which normally presents as a cloud-like dispersion of the image that do not correspond to any physically meaningful distortion. This must be taken into account when making further adjustments to the position or orientation of the registered histology image in MRI space, e.g. by appending and optimising new transformations as opposed to changing the existing ones.

Once the combined chains were created, the histology images were loaded at the desired resolution (8 µm/px) from the the original SVS files, and the MRI volumes from the NIfTI files with their respective *sform* matrix. The combined chain was assigned to the domain of the histology image, and the MR volume was evaluated at the transformed histology domain to obtain a 2-D histology-equivalent MR image for each histology slide.

6.1.3 Results

Stage 1 registrations

Figure 6.3A shows a representative stage-1 registration result. In most cases the registrations were visually accurate and all deformations appeared physically realistic. In a limited number of cases (11 out of 88) the default configurations had to be manually adjusted to achieve satisfactory alignment. Examples of these are shown in panels B and C of Figure 6.3, and are discussed below.

The most commonly observed pattern of misregistration was the stretching of the histology image beyond the edges of the cut surface. In these cases either a

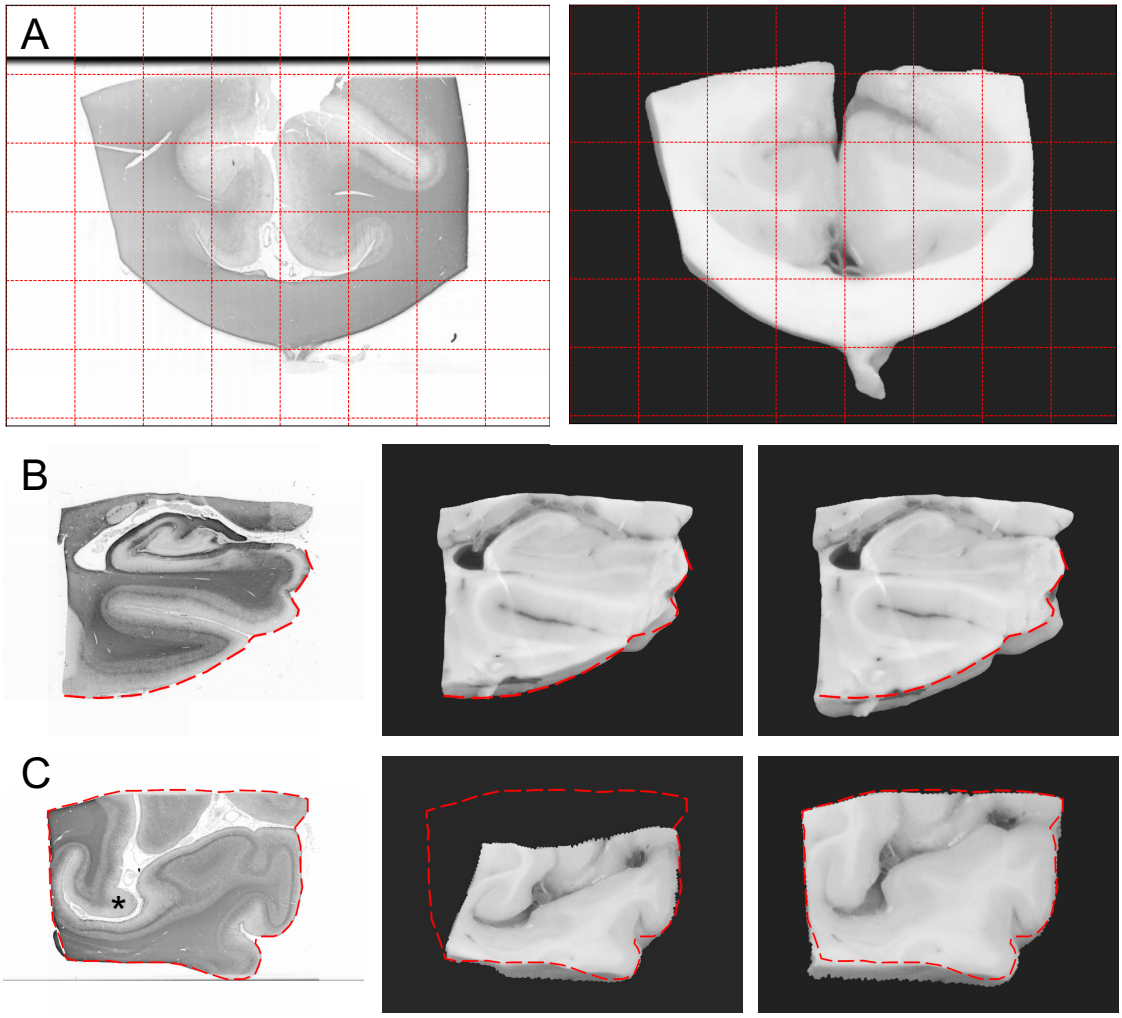

Figure 6.3: Stage-1 registration results. **(A)** Representative stage-1 registration result shown by the aligned histology (*left*) and tissue block (*right*) images of a corpus callosum specimen. Grid spacing: 5 mm. **(B)** Misregistration by stretching the histology section over the side surface of the block. Confining the mask to the cut surface leads to accurate registration. **(C)** Incorrectly specified image resolutions lead to erroneous scaling. Changing the resolution input fixes the problem. The asterisk highlights a slight anatomical discrepancy between the images.

side of the block was visible because the photograph was taken at an angle, or a neighbouring gyrus was visible despite the block being centred at the optical axis. Both of these problems were successfully addressed using binary masks that were created manually for the tissue block photographs to exclude non-matching regions from the registration. A corresponding example is shown in Figure 6.3B.

In other cases, the images were aligned for some major image features, but not all, and the tissue block photograph appeared to be too small compared to the

histology image. In these cases the resolution of the tissue block photograph was re-verified, and found to be largely different from the original estimate, because the position of the camera had been changed before the tissue blocks were photographed. The images registered properly after the resolution of the tissue block resolution was decreased (from 50 µm/px to 70 µm/px). A corresponding example is shown in Figure 6.3C.

Occasionally, minor anatomical discrepancies were observed between the histology image and the tissue block, leading to an apparent misalignment of the images in these regions. The reason why such discrepancies arise is in the sampling of the tissue for histology. The specific section that is being registered may come from deeper than a millimetre beneath the surface of the block, after several other slices have been peeled in the microtome for other stains or for quality reasons. At such depths, the anatomy might be visibly different from what is shown on the surface of the block. For the sections with anatomical discrepancies a higher regularisation weight of 0.8 was uniformly employed as a compromise solution. This is to compensate for bulk deformations, but also to minimise artificial distortions that are driven by the mismatch of the anatomical pattern at smaller length scales. An example of a typical discrepancy is marked with an asterisk in Figure 6.3C.

Stage 2 registrations

A representative result of the stage-2 registrations is displayed on a specimen from the *visual cortex* in Figure 6.4 (panels A and B). In nearly all 88 cases, the alignment of the tissue contours and even the perforating vessels were observed as highly accurate, which is in line with the ≤ 0.2 mm registration error that was measured on *hippocampal* and *callosal* blocks in Chapter 4. Larger misalignments (≤ 1 mm) were occasionally (N=3) found with blocks from the *anterior cingulate* region, where the two opposite facing gyri had undergone relatively large elastic deformations, closing or opening the interhemispheric fissure. An example of this can be seen on panels C-E in Figure 6.4.

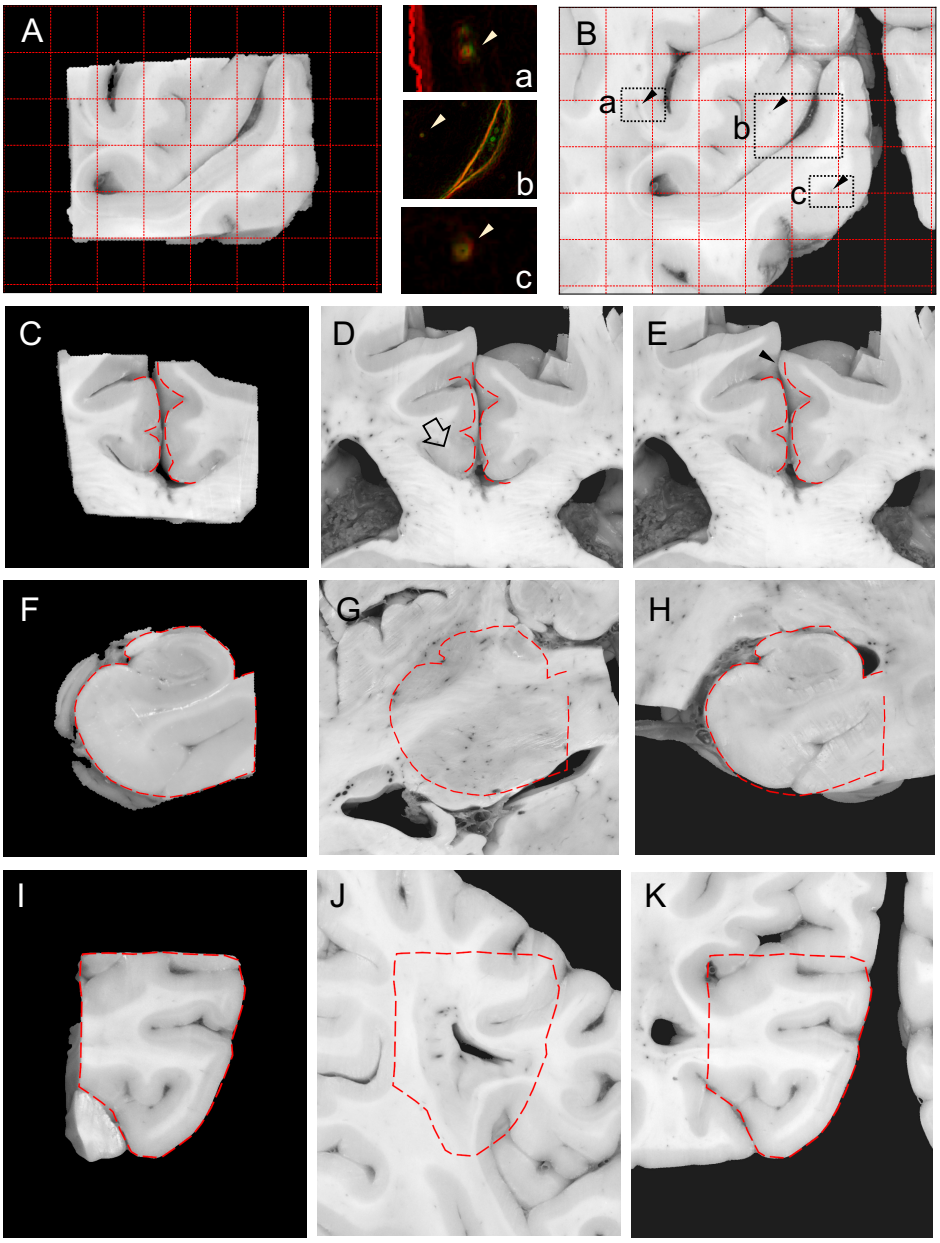

Figure 6.4: Stage-2 registration results. **(A-B)** Representative stage-2 registration result, shown by the alignment between the tissue block and the corresponding portion of the slice photograph for a visual cortical specimen. Grid spacing: 5 mm. **(a,b,c)** Fused red and green images showing close correspondence between perforating vessels (*arrowheads*) in the registered photographs. **(C-E)** Deflections of the cingulate gyri (*large arrow*) are responsible for the largest observed misalignments in stage 2 (*arrowhead*). **(F-H)** Misidentification of the insertion site for a hippocampal block (G), where the location of the insertion site was offset from its true location by 14.7 mm. Manual specification lead to identical high-quality registration (H). **(I-K)** Misidentification of the insertion site for a visual cortical specimen, where confounding anatomical patterns led to the incorrect estimation of the block rotation (J). Refining the rotation search led to a high-quality registration (K).

The default parameters had to be manually adjusted to achieve similar high-quality registrations in a total of 6 cases. In five of these, the automatically determined insertion site was too far from the actual centre of the block, therefore the correct alignment could not be reached within the default parameter bounds, and the block was registered to another insertion site. These issues were fully resolved by providing the insertion site coordinates manually. An example of these is show in panels F-H in Figure 6.4. In the remaining one case, despite the coordinates being given in reasonable distance, the insertion site was still incorrectly identified. In this case refining the rotation search from 30° increments to 15° increments led to the resolution of the problem as well. The case is shown in panels I-K in Figure 6.4.

Stage 3 registrations

In the majority (N=161) of the 209 cases the registrations were visually accurate to an extent that even perforating vessels (smaller than a millimetre) were closely aligned between the MRI and the brain slice photograph. However, in several cases this level of accuracy was specific to one of the hemispheres and not necessarily the other. A representative example of this phenomenon is shown in Figure 6.5 (panels A and B). Such observations were more accented on posterior slices, where the hemispheres were no longer connected by the corpus callosum and posterior commissure, hence they could move independently. After careful observation of the problematic slices, three major types of distortions were identified that could not always be satisfactorily compensated for by stage-3 registration: (1) the closing or extreme opening of the interhemispheric fissure due to the positioning of the hemispheres for the photograph, (2) extreme rotations of one hemisphere relative to the other during the MR acquisition, and (3) different slicing position and/or orientation for the left and right hemispheres, that could have been accommodated only if a sharp bend had been introduced to the photograph during the 4th substage of the registration.

To address these issues, the registration of the slices was repeated twice, separately for each hemisphere with appropriate cost function masking. Anatomical

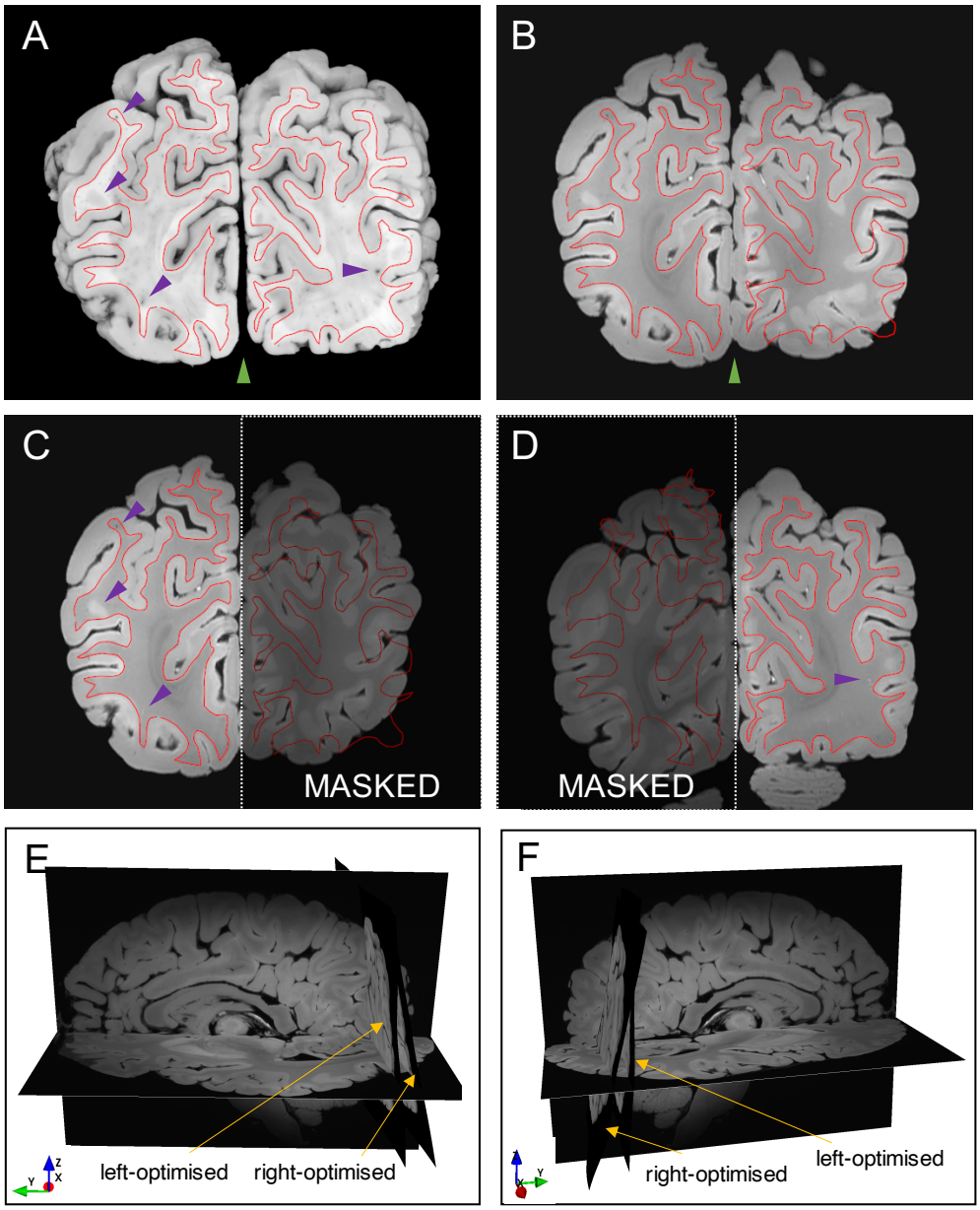

Figure 6.5: Stage-3 registration results. All images are shown in neurological orientation. **(A)** Tissue block photograph with highlighted gray-white matter contour. The *purple arrowheads* point at perforating vessels to assess the accuracy of the registration of the slice with the MRI volume. The *green arrowhead* shows the artificial opening of the interhemispheric fissure that leads to misregistration. **(B)** Registered 2-D MRI slice obtained with whole-slice masking. The gray-white matter contour is overlaid from the slice photograph, indicating highly accurate registration in the left hemisphere, and substantial misalignments towards the bottom of the right hemisphere. **(C)** 2-D registered MRI slice obtained with left hemisphere masking. **(D)** 2-D registered MRI slice obtained with right hemisphere masking. Both unilaterally-optimised registrations are highly accurate, as indicated by the alignment of the contours and perforating vessels. **(E-F)** The 3-D rendering of the unilaterally registered slices reveals the large difference in the slicing position and orientation between the left and the right hemispheres.

labels that were generated by the BrainSuite Cortical Surface Extraction pipeline based on the structural MRI volume were fused to create separate binary masks for the left and the right hemispheres (excluding the cerebellum). The binary masks were then specified in the stage-3 configurations as masks for the MR volume. In 5 of the 14 subjects, the BrainSuite pipeline failed to create accurate anatomical segmentations (the BrainSuite pipeline is not optimised for post-mortem scans). In these subjects hemisphere-specific binary masks were created manually for each coronal section, which led to identical high-quality results. A representative example of the hemisphere-selective stage-3 registration results is shown in Figure 6.5 (panels C and D). Panels E and F illustrate the large differences that we occasionally encountered in the slicing position and slicing orientation of the hemispheres.

Histology-MRI registrations in the sample brain

As we have seen, some amount of manual intervention was necessary when the pipeline was applied to the entire dataset, but eventually all of these led to successful registrations in all three stages, and the number of interventions can be significantly reduced by implementing changes at the data collection stage based on the above experiences.

After the stage-specific registrations, we had a seamless experience of combining the optimised transformation chains from stages 1-3 and resampling the MRI data on the transformed histology domains at 8 µm/px resolution. First we show the results from the sample brain, where altogether 12 slides were registered from different anatomical regions.

Figure 6.6 shows the voxelwise correspondent histology and MR images for those specimens that belong to a single hemisphere only, and the thalamus block, which is bilateral. With the exception of the thalamus block, all the histology images had similar sizes, and were stained for PLP. The *thalamus* section was almost twice as large as the other slides, and it was stained for Ferritin, which gave a similar gray-white matter contrast as PLP. The reason why the thalamus section is shown with the other unilateral blocks is that the resampled MR images for this section

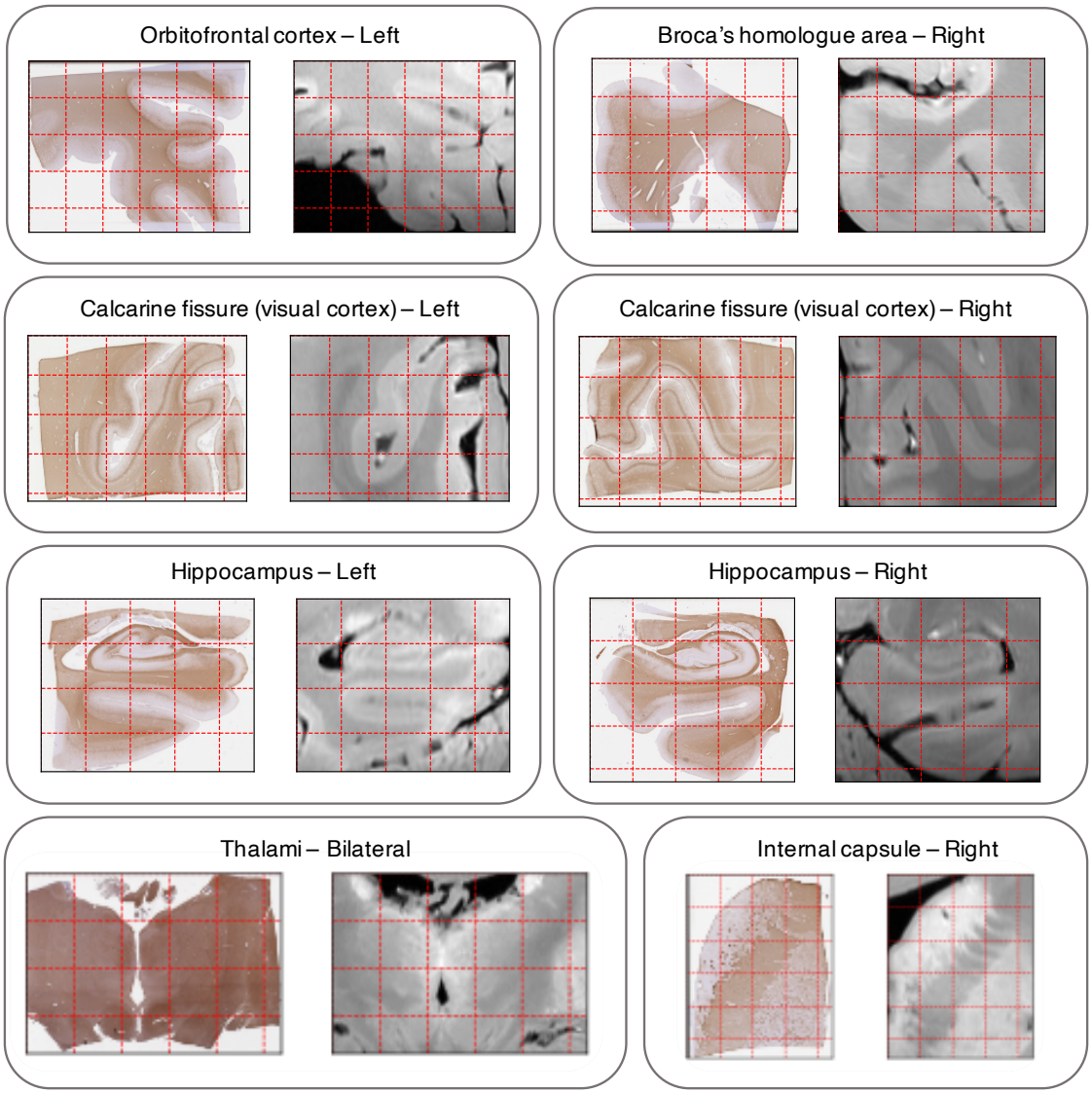

Figure 6.6: Histology-to-MRI registration results for various anatomical regions in the sample brain. Grid spacing: 10 mm for the thalamus, 5 mm otherwise. These results were obtained with a unilaterally optimised stage-3 registration, except for the thalamus section, where stage 3 was optimised bilaterally. The histology sections were immunostained for PLP, except for the thalamus that was inmmunostained for Ferritin. The MR images were resampled from the 0.25 mm structural (TRUFI) volume.

were visually indistinguishable regardless whether the bilateral, left-optimised, or right-optimised stage-3 transformation chain was used. This is most likely explained by the restricted motion of the hemispheres within this block.

In the blocks from the *corpus callosum* and the *anterior cingulate* regions, the gyri from the two hemispheres had much greater freedom to move independently. Consequently, in these regions we did observe differences between the MR images

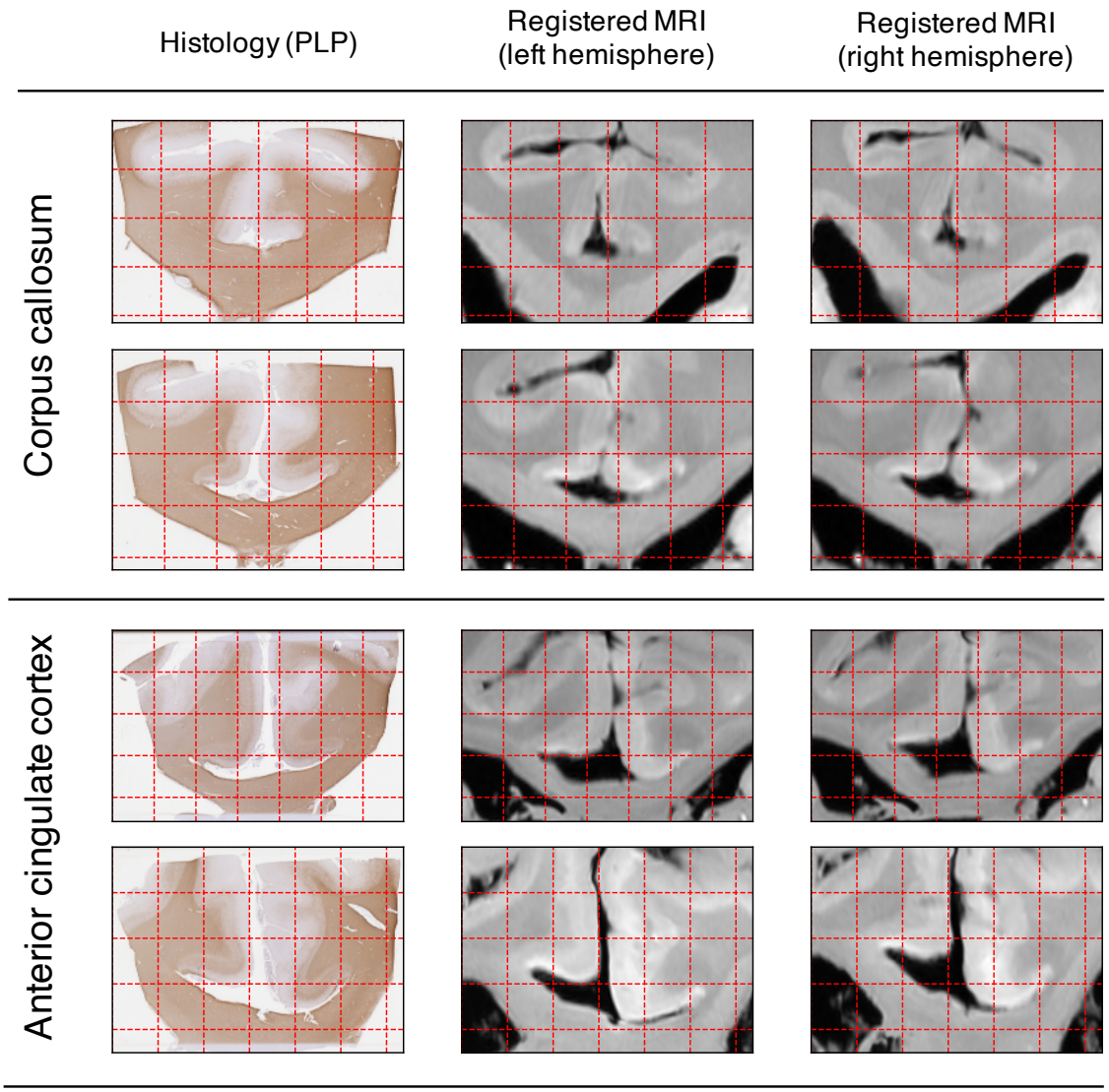

Figure 6.7: A visual comparison of the histology-to-MRI registrations on the 4 central blocks of the sample brain obtained with left-optimised and right-optimised stage-3 registration. Grid spacing: 5 mm. The histology sections were immunostained for PLP, and the MR images were resampled from the 0.25 mm structural (TRUFI) volume.

that were reconstructed using the left-optimised and the right-optimised stage-3 transformation chains. A comparison of the two reconstructions is shown in Figure 6.7 for the 4 blocks that were available from the same subject. While these images represent good alignment overall, some larger differences were observed at the interface of the cingulate gyri and the corpus callosum. In these regions both the left-optimised and the right-optimised images exhibited a mutually more precise alignment on the side of the optimisation, and this was normally more accurate

than what could be achieved with bilateral optimisation. The recommendation is therefore to use unilaterally optimised registrations of the callosal and cingulate blocks whenever ROIs can be separately considered for each hemisphere, and use the bilateral result only when the ROIs span across both hemispheres.

Histology-MRI registrations in other brains

A standard set of 6 histology sections was registered for all 14 brains to test the robustness of the combined pipeline as described above. The following figures (Figure 6.8-6.14) display registration results from different subjects, and overall they provide a representative set of results for all tested anatomical regions in the entire cohort. The histology and MR images are presented at 8 µm resolution. The selected subjects include both diseased and control brains, and individuals from both ends of the age spectrum (see details of the subjects in the captions). The visual cortex and hippocampal specimens are presented from the right hemisphere for consistency, and the respective MR images were generated using right-optimised stage-3 transformation chains. The MR images for the callosal and cingulate specimens were uniformly generated using bilaterally optimised stage-3 transformation chains.

As it can be seen in the figures, the combined registration pipeline provides a mapping between histology and MR data that is accurate on the submillimetre scale, and is consistent across several anatomical regions and subjects. The accuracy was found to be the highest on the feature-rich specimens of the visual cortex and the hippocampus, with occasional misalignments that were larger but still submillimetre for the cingulate gyri in the other two regions.

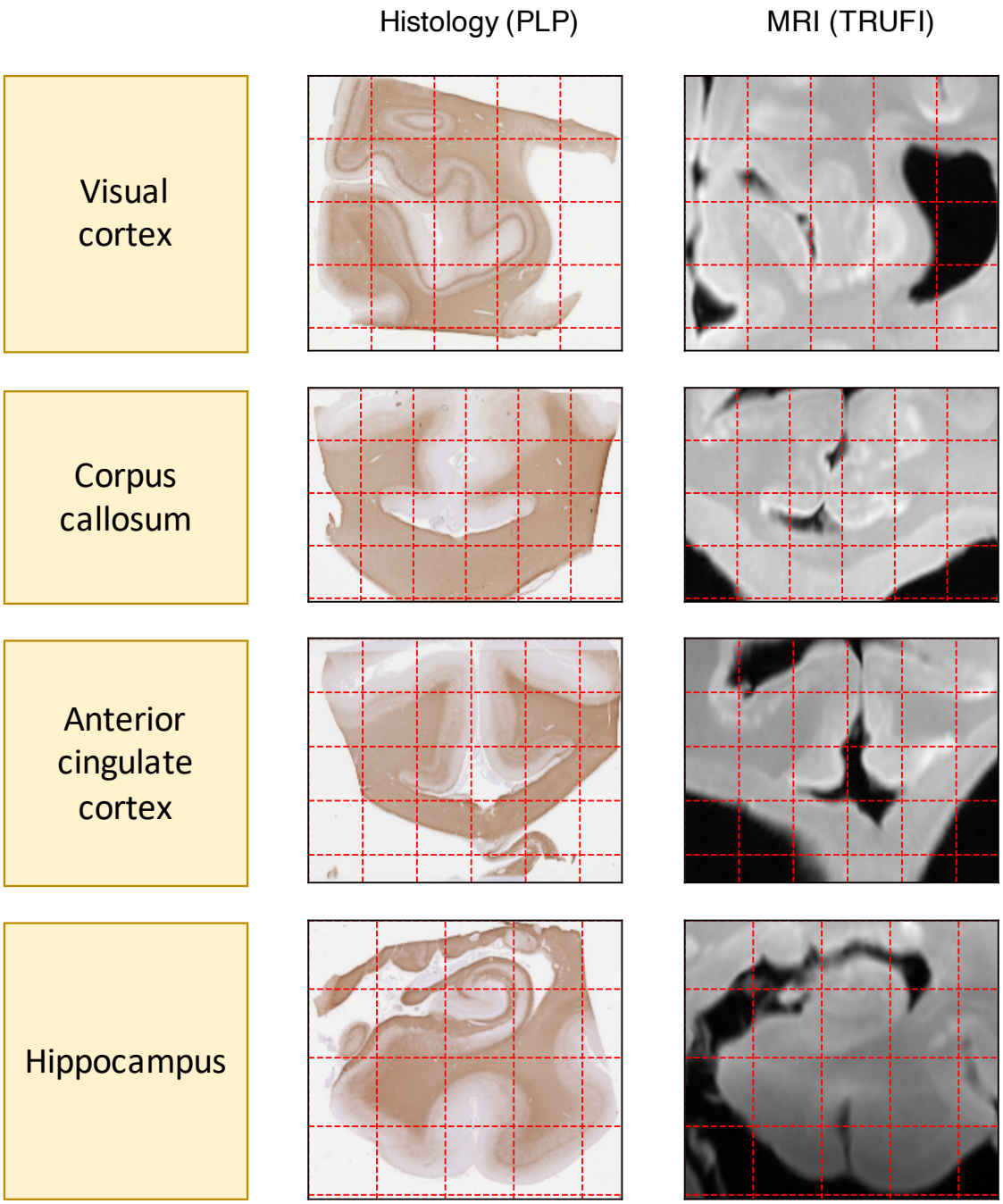

Figure 6.8: Representative histology-to-MRI registrations for each anatomical region in subject NP004-15. Age: 65 years, Sex: male, Diagnosis: ALS.

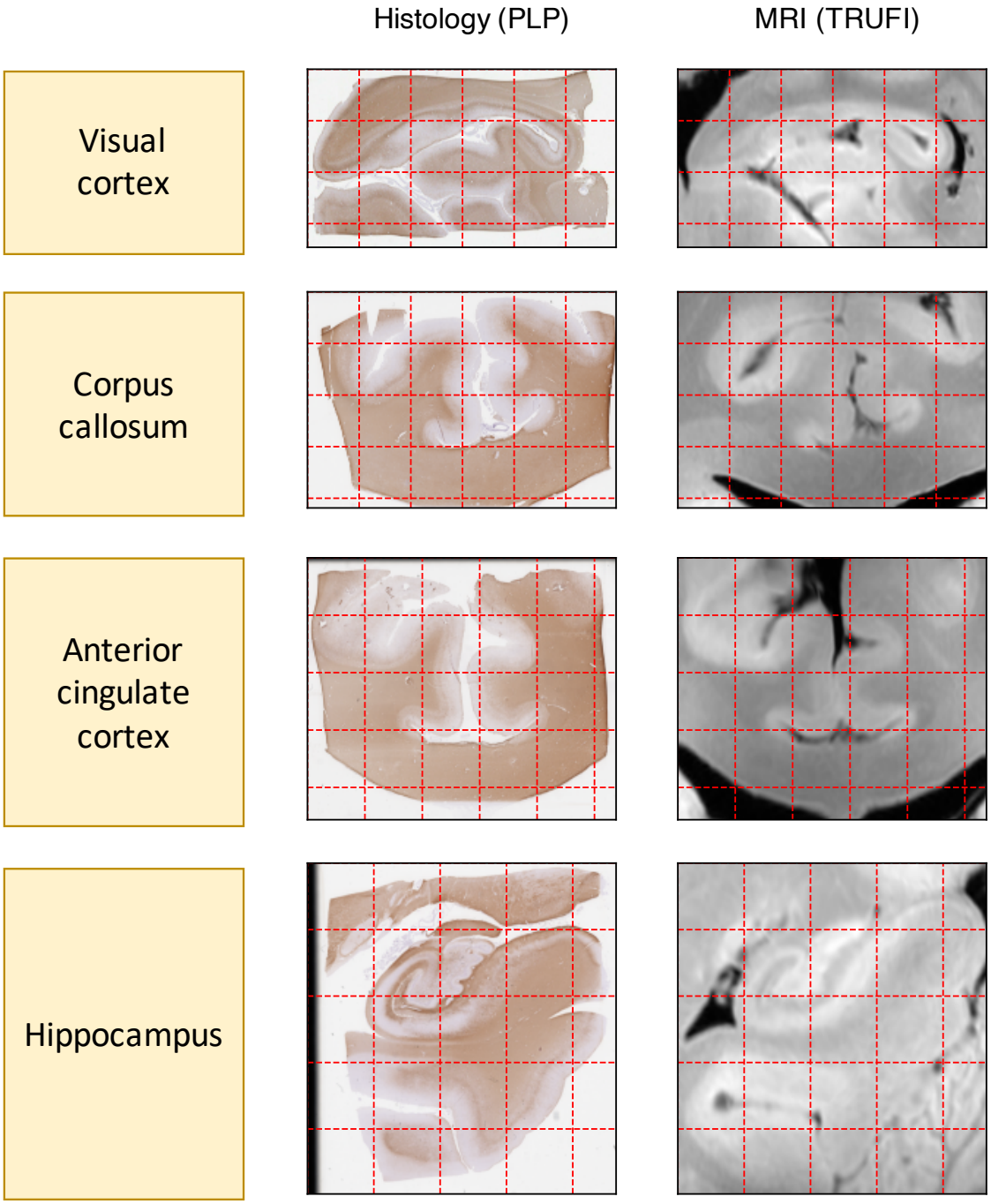

Figure 6.9: Representative histology-to-MRI registrations for each anatomical region in subject NP021-15. Age: 65 years, Sex: male, Diagnosis: uncertain.

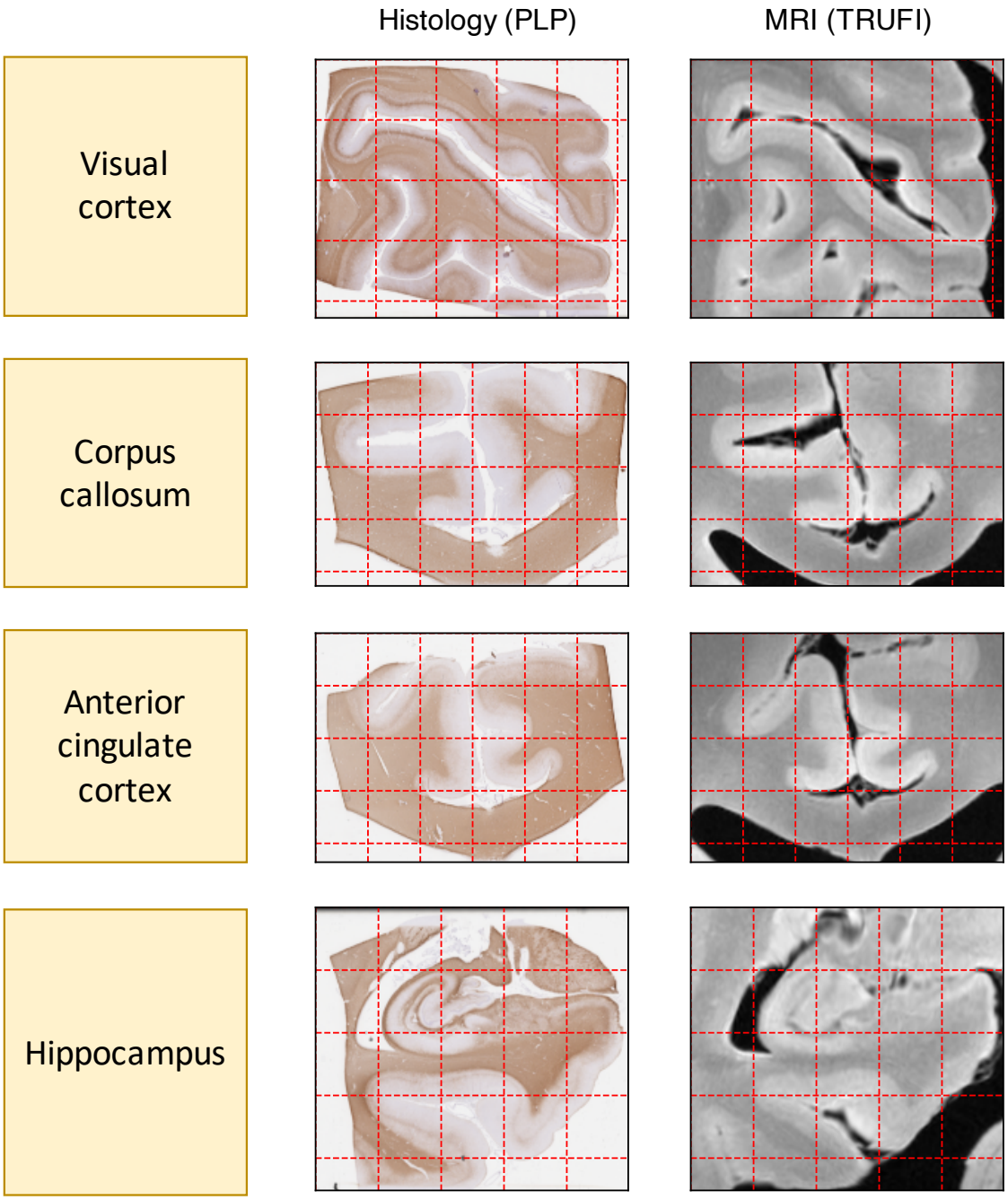

Figure 6.10: Representative histology-to-MRI registrations for each anatomical region in subject NP029-16. Age: 79 years, Sex: female, Diagnosis: ALS.

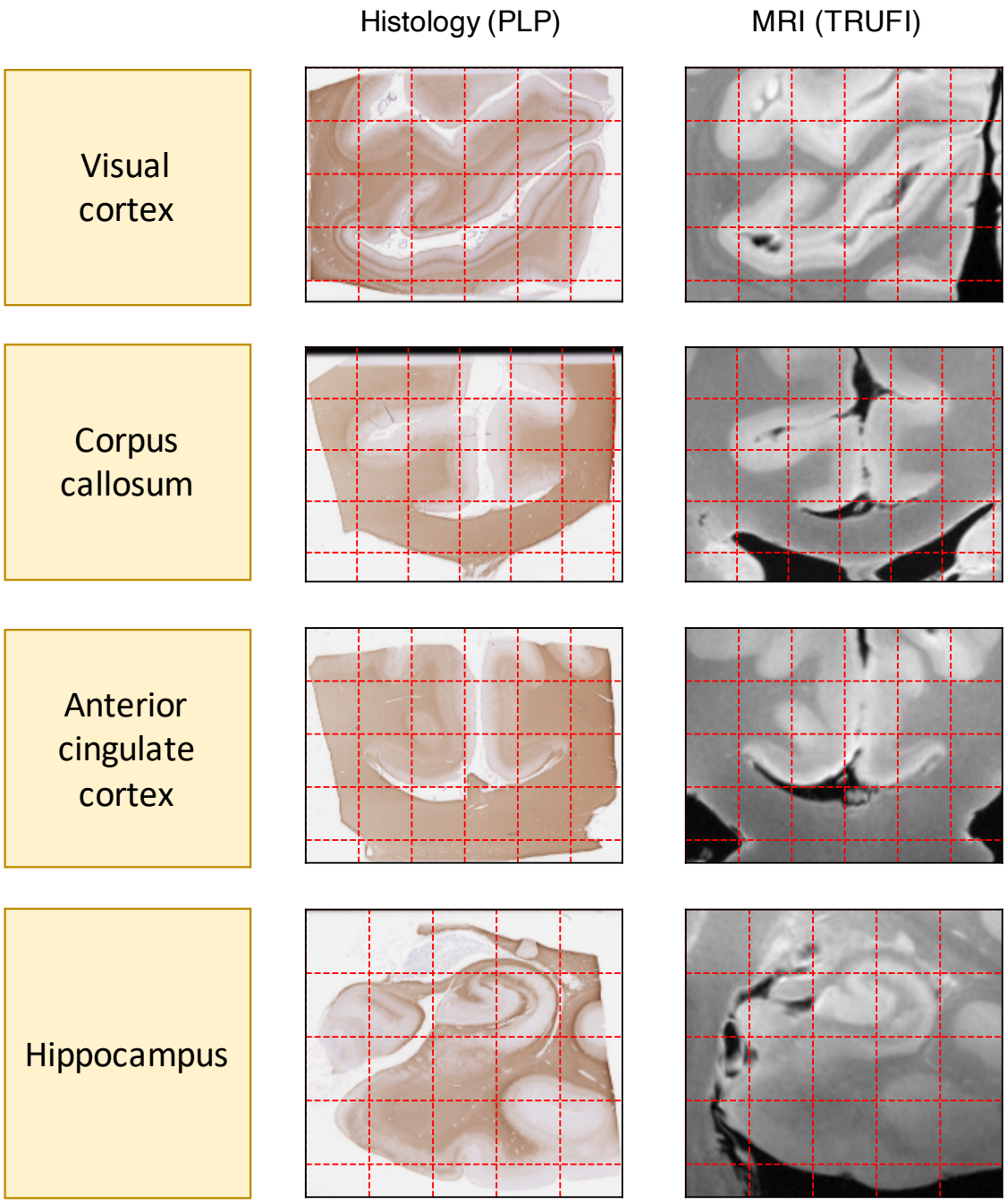

Figure 6.11: Representative histology-to-MRI registrations for each anatomical region in subject NP040-16. Age: 61 years, Sex: male, Diagnosis: no neuropathology.

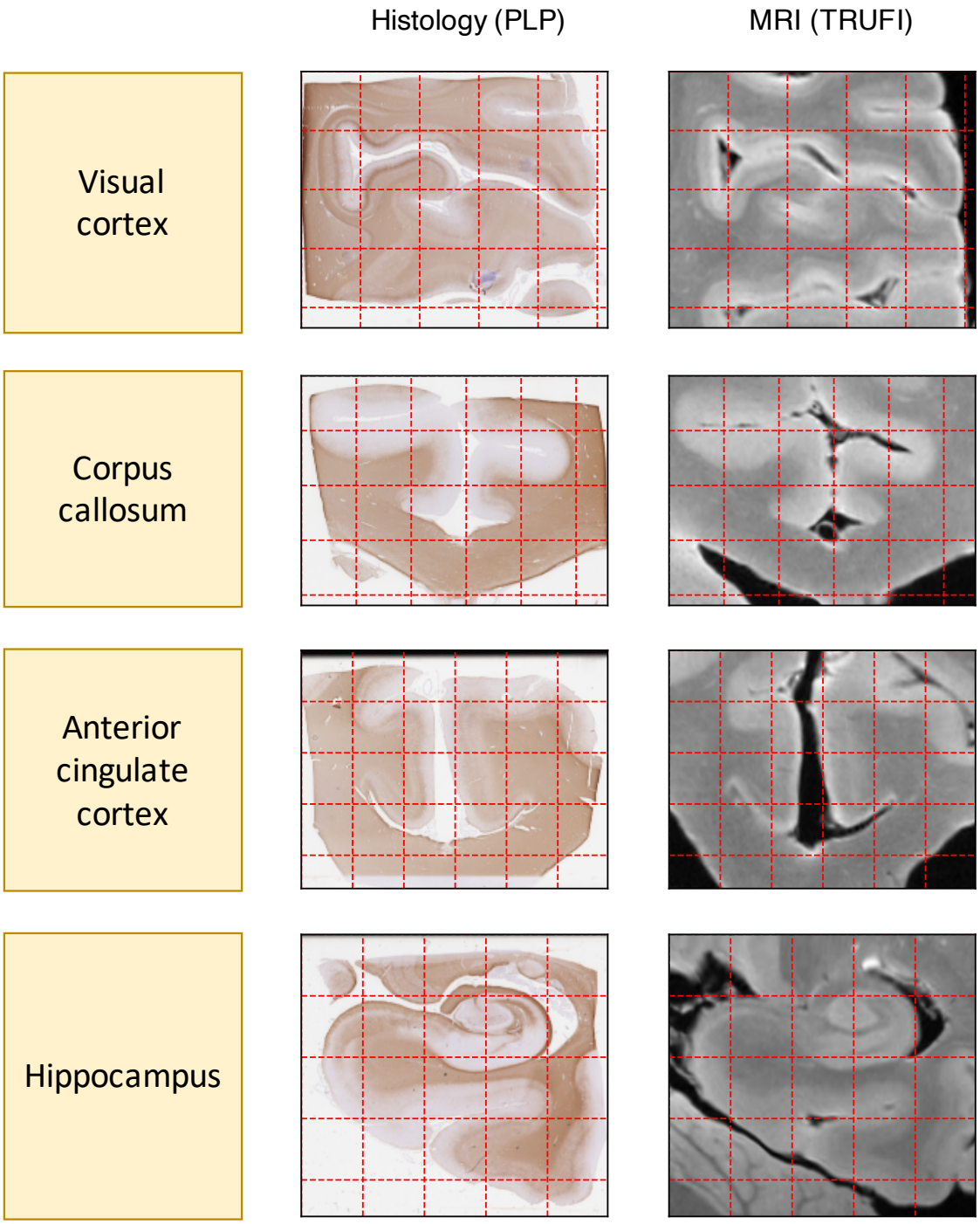

Figure 6.12: Representative histology-to-MRI registrations for each anatomical region in subject NP096-15. Age: 71 years, Sex: male, Diagnosis: ALS.

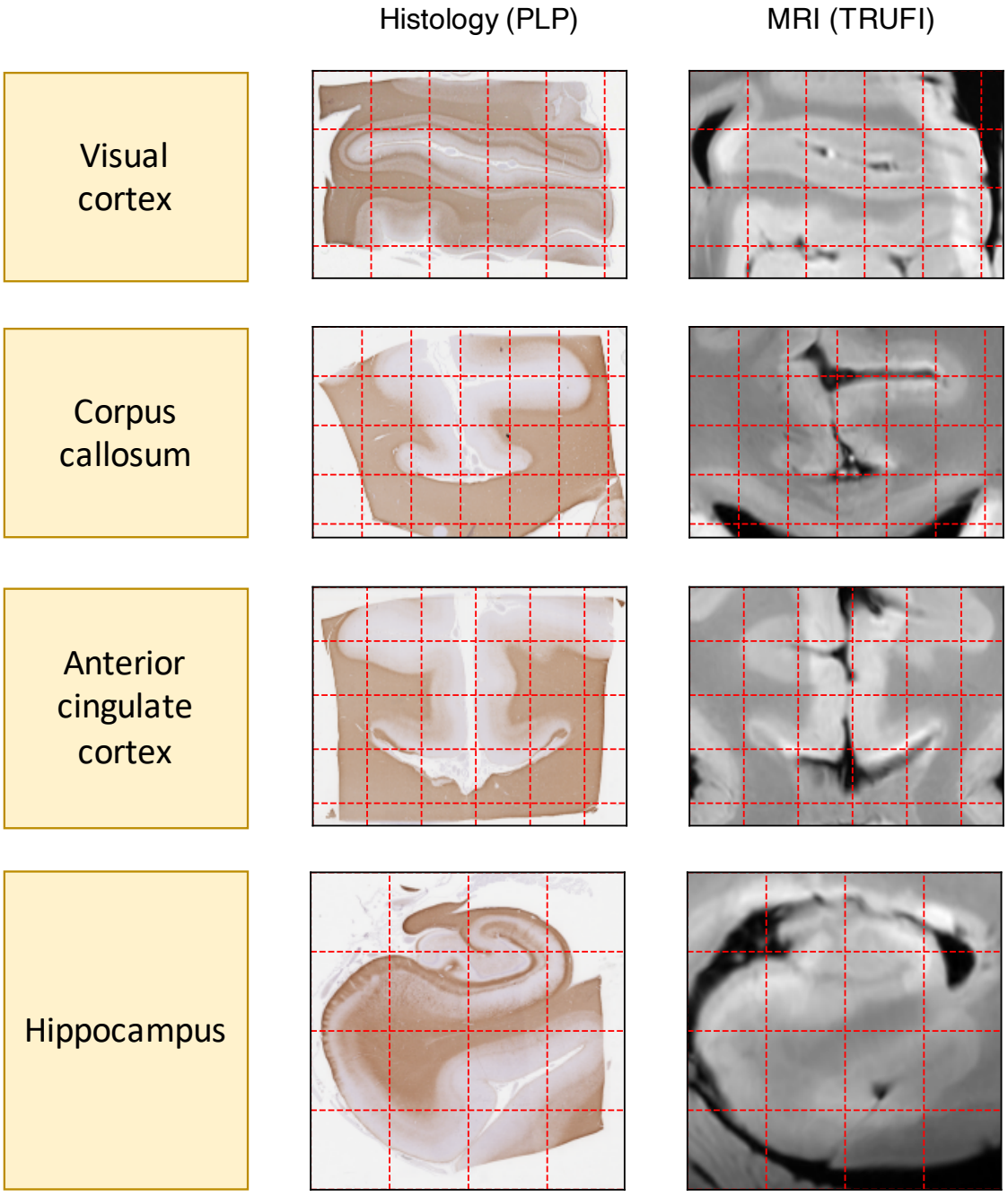

Figure 6.13: Representative histology-to-MRI registrations for each anatomical region in subject NP111-14. Age: 27 years, Sex: male, Diagnosis: ALS.

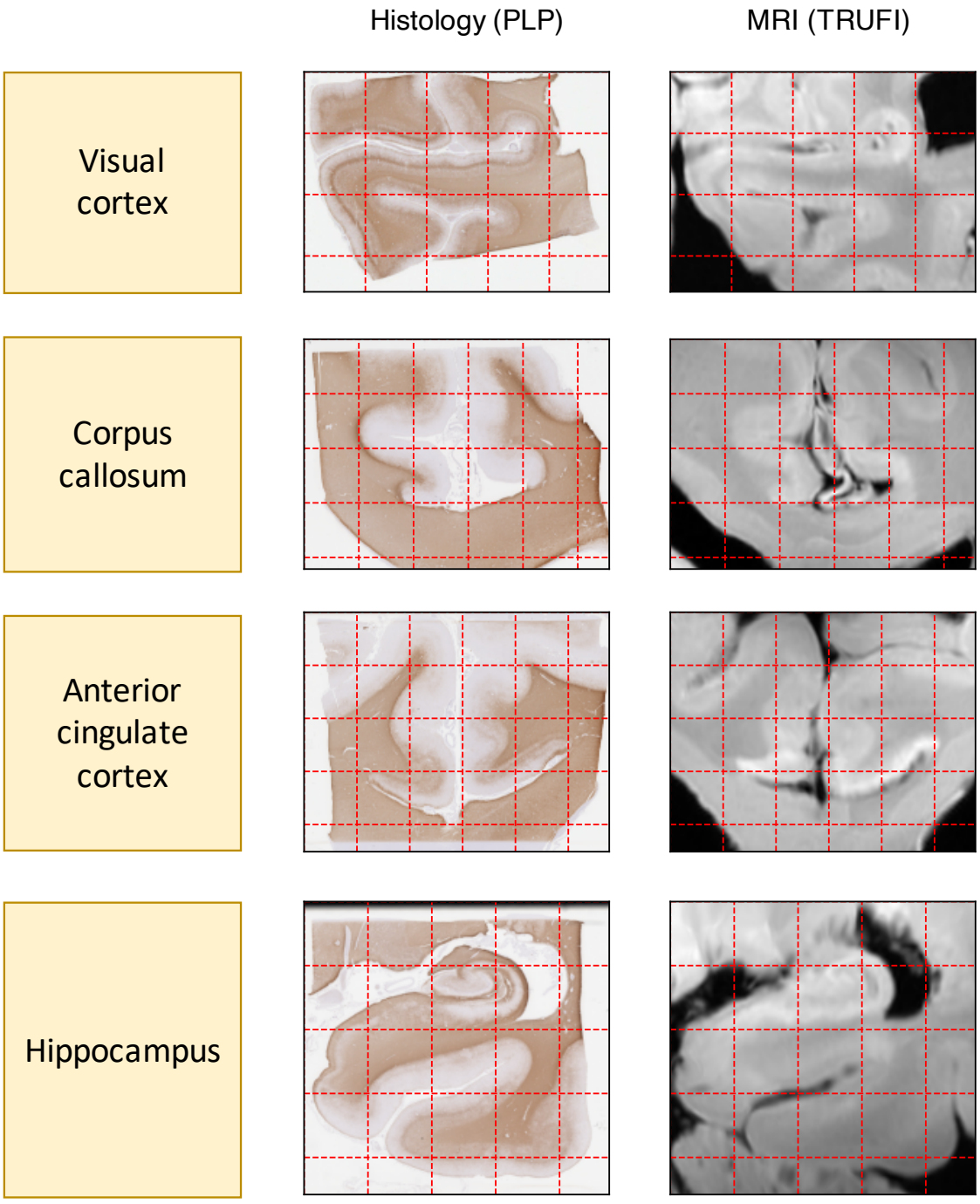

Figure 6.14: Representative histology-to-MRI registrations for each anatomical region in subject NP135-14. Age: 76 years, Sex: male, Diagnosis: ALS.

Stage 4: post-hoc optimisation

While discussing stage-1 configurations, we briefly mentioned that the histology image and the tissue block photograph may become anatomically inconsistent if the block is sampled at a reasonable depth from its original photographed surface. This could be the result of acquiring an excessive number of sections from the same block (possibly for diagnostic or quality control reasons). In this case the sampling depth would be solely responsible for the change in the anatomy. However the inconsistencies can sometimes be seen even when only a handful of sections are acquired. Considering that individual sections are normally no thicker than 10 µm, and the anatomy is usually consistent over these length scales, a more likely explanation of the inconsistencies is shown by the schematic in Figure 6.15A.

As the blocks are sampled by hand, a certain degree of unevenness or non-parallelism of the block surfaces is unavoidable. When the block is embedded in paraffin, the upper surface is therefore not perfectly horizontal. When sectioning begins in the microtome, the embedded block must be trimmed down to the point where the whole surface area of the tissue block is evenly exposed. Depending on the tilt of upper surface, the first full slice of tissue may come from as deep as 0.5 – 1 mm (corresponding to an angle of 2° for a 30 mm long block). In the case of multiple stains, or when stains need to be repeated for quality reasons, this problem is further exaggerated: the deeper the block is sampled, the less consistent the stained histological sections will be with the surface anatomy of the blocks as seen in the photographs.

To minimise the effect of such inconsistencies on the accuracy of the histology-MRI registration, a higher regularisation weight was used in stage 1 to preserve the structural self-consistency of the histological section while compensating for bulk deformations. In this section, we explore the feasibility of estimating the slicing depth and angle by extending the pipeline with an optional 4th stage that performs a rigid slice-to-volume registration of the already registered histology section in MRI space.

In stage 4 the combined transformation chain is attached to the histology image, and extended by two transformations: a 3-D rotation (about the centre of the

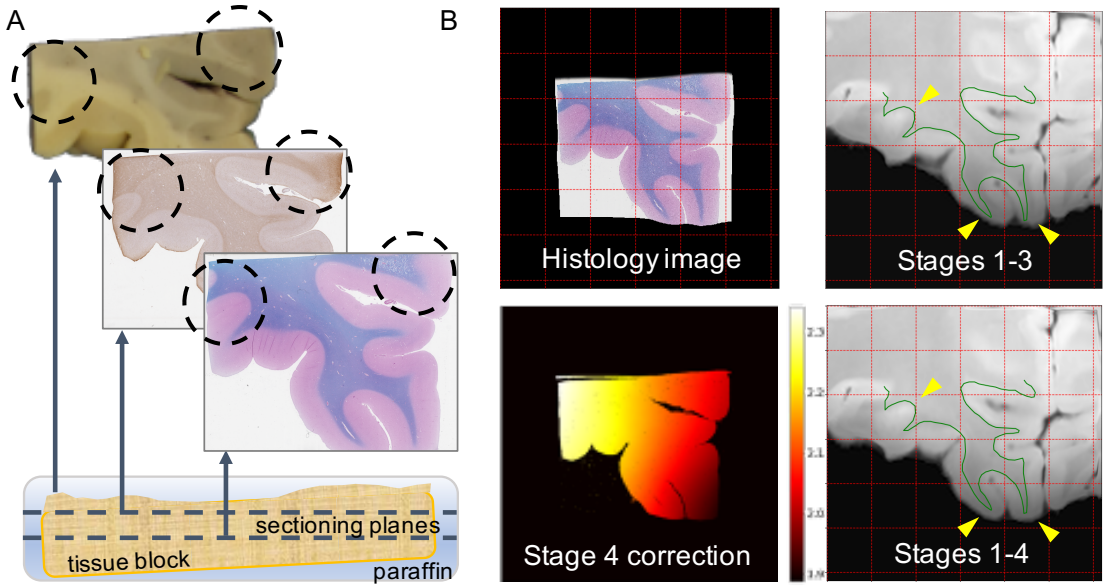

Figure 6.15: (A) The source of anatomical inconsistencies between consecutive histology sections and the tissue block photograph (see the main text). (B) Application of the optional 4th stage to estimate the slicing depth and slicing angle. The colour map reveals significant pixelwise distances between the adjusted section and the histology section as it was inserted by stages 1-3. Grid spacing: 5 mm.

registered histology image) and a 3-D translation. It is crucial that stage 4 only optimises the parameters of these two transformations, and treats all parameters in the combined transformation chain as constant, otherwise it would disrupt the integrity of the chain. The execution of stage 4 relies on the stage-3 routine, but stage 4 is different in several key aspects:

1. The 2-D image input is the histology image, not the brain slice. As MR data is likely to be present outside the edges of the histology section, the bright background of the image is masked out by thresholding the grayscale image at the 90% intensity level.

2. Only substage 1 is executed to provide a rigid compensation, and the affine and non-linear substages are omitted to prevent the optimisation from overfitting to artefacts of the histology image when higher degrees of freedom are used.

3. The grid search of substage 1 takes altogether 81 initialisations at 1/50th scale (0.4 mm/px). First the centre and the normal vector of the registered

histology image are calculated as the mean and the 3rd principal component of the coordinates of the histology image in MRI space. Then the histology section is initialised at 3 equidistant points: $\begin{bmatrix}-2 & 0 & 2\end{bmatrix}$ mm away from the centre in the direction of the normal vector. At each position, 3 values are tested for each of the three rotation components: $\begin{bmatrix}-5^\circ & 0 & 5^\circ\end{bmatrix}$.

4. Starting from each initialisation, the 3-D rotation and 3-D translation parameters are fine-tuned in a single course of BOBYQA optimisation to minimise the MIND cost (computed after the resampling of the MR volume, for a 3×3 dense neighbourhood, $\sigma = 1$). Contrary to the stage 3 routine, the scale parameters and the 2-D translation parameters are not affected during the execution of substage 1.

Figure 6.15B shows the effect of applying the additional 4th stage on a section of the orbitofrontal cortex. The histology image used in this experiment was a late addition to an existing set of images, and it was stained with LFB-PAS to maximise the gray-white matter contrast. Slight anatomical inconsistencies were already visible in some of the early sections, but they became more exaggerated on the latest one. Despite exhibiting a strong gray-white matter contrast, the LFB-PAS-stained section was visibly misaligned with the MR data after it was registered with stages 1-3. Accordingly, stage 4 identified an approximately 2-mm shift for the section (slicing depth) with a diagonal tilt corresponding to a 0.4-mm elevation difference between opposing sides of the section (equivalent to an approximate slicing angle of 2°). After stage 4, the gray-white matter contour was well aligned with the MR data for the entire section.

The successful application of stage 4 in the above case provides further evidence for the precision of the TIRL slice-to-volume registration routine. It is also an important demonstration that the misalignments related to sectioning differences can be mitigated by an optional post-hoc optimisation, which allows accurate registration for those sections that might be sampled from the same blocks in the future.

6.1.4 Discussion

Based on these results, there seem to be several key advantages of using the TIRL pipeline for the joint analysis of histology and MR data as opposed to the common practice of manually defining ROIs:

1. The TIRL pipeline takes advantage of the whole brain slice to provide a more accurate localisation. On the contrary, when ROIs are defined by hand, it is often hard or impossible to perceive depth and orientation with such precision, especially if the appearance of the anatomical region is similar on several consecutive MR slices.

2. On the majority of the sections, where the TIRL pipeline achieves sufficient accuracy, the MR signal can be directly related to the underlying microstructural information in a voxelwise fashion. Manual ROIs on the other hand can only provide regional averages.

3. Beyond scalar metrics, MRI and histology-derived vector and tensor quantities are also possible to compare, as they can be adjusted for local deformations given the optimised transformation chain. Such adjustments are not possible when ROIs are defined manually.

4. The registration process is largely automated, therefore it requires less manual effort and leads to more reproducible results.

Aligning histology with multi-modal MR data

The inspiration for the post-mortem ALS study and the development of TIRL is to ultimately compare multi-modal MR data with multi-modal ground-truth histological data in key anatomical regions. An example of such a comparison in the visual cortex is shown in Figure 6.16 between PLP histology data and diffusion tensor derived scalar metrics such as mean diffusivity (MD) and fractional anisotropy (FA). The latter two were resampled from output volumes of FSL DTIFIT, that

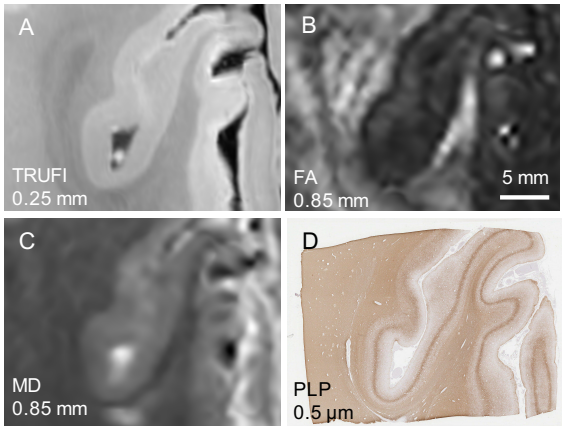

Figure 6.16: Relating histology to multi-modal MRI data.

were in register with the TRUFI volume that was used to estimate the parameters of the combined transformation chain.

When the MR modalities are not in the same space, the transformations between them should be first estimated with FSL FLIRT [99], and the respective FLIRT matrices can be readily appended to the combined TIRL chain to achieve the desired mapping without multiple interpolations. If the histology image needs to be transformed to the space of an atlas, FNIRT warp fields are equally simple to incorporate in TIRL chains.

Furthermore, TIRL implements the functionality of FSL *vecreg* to rotate vectors and tensors through any transformation chain (also including FLIRT and FNIRT transformations) according to the following two paradigms [100]: *finite strain* (i.e. rotation-only) paradigm, and the *preservation of principal eigenvectors* paradigm. These are essential new developments that will in the future facilitate the comparison of directional information (e.g. primary fibre orientations) between histology and MR modalities.

6.2 Registration by one photo intermediate

6.2.1 Dataset

The dissection protocol for the post-mortem ALS dataset (described in Section 3.1) distinguished the motor cortex from all other parts of the brain by employing

a different sampling method for the tissue blocks. To preserve the integrity of the transition zone between the primary motor (M1) and primary sensory (S1) cortices at the bottom of the central sulcus, the precentral and postcentral gyri were sampled together. Before the coronal slicing of the brain, 1 cm thick rectangular blocks were removed from the somatotopic representation areas of the leg (most medial), the hand (centre) and the face (most lateral). Figure 6.17 shows the sampling of such an M1S1 block from the hand area in the right hemisphere. The corresponding histology section was taken from the middle of the block, for which the block was cut in half, and the intermediate surface was photographed to guide the registration of the histology image.

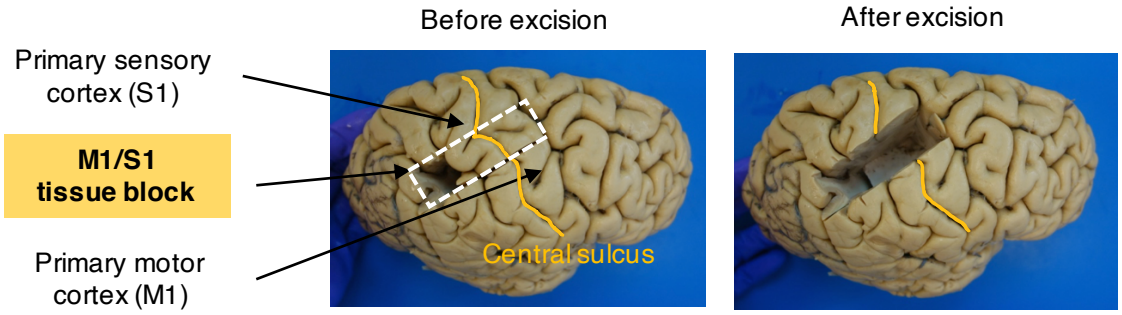

Figure 6.17: Sampling an M1S1 block from the right hand area.

Due to the special sampling method, the 3-stage pipeline that was successfully applied to register the extramotor specimens, is not directly applicable in the case of the M1S1 blocks. In this section, we show how the registration pipeline can be adapted to register a single histological image from the hand area of the left hemisphere, achieving histology-to-MRI registration with a single photographic intermediate.

6.2.2 Methods

Since there is only a single photographic intermediate (the tissue block photograph), stage 2 is omitted, leaving the M1S1 pipeline with the following two stages: (1) histology-to-block registration (identical to stage 1), and (2) block-to-volume registration (equivalent to stage 3 with modifications). Algorithmic and implementation details for the underlying registration routines were described in Chapters 3 and 5, respectively. For this demonstration, the *histology-to-block registration* was

executed with identical parameter configurations as previously for the extramotor specimens (see Section 6.1.2).

In the *block-to-volume stage*, the background-segmented tissue block photograph is registered to the structural (TRUFI) MR volume at 0.25 mm isotropic resolution. As with the coronal slices, this 2D-to-3D registration also relies on reasonable initial estimates of both the position and orientation of the tissue block image. These are far from standard and have to be determined manually, because the sampling of the blocks is guided by the gyral patterns of the intact brain surface. Consequently, it is also exceptionally hard to provide the necessary initialisation details while looking at the MR volume in orthogonal slice view. To make the initialisation process suitably accurate for block-to-volume registration, the stage-3 registration of the coronal slices is exploited instead of relying on the gyral patterns.

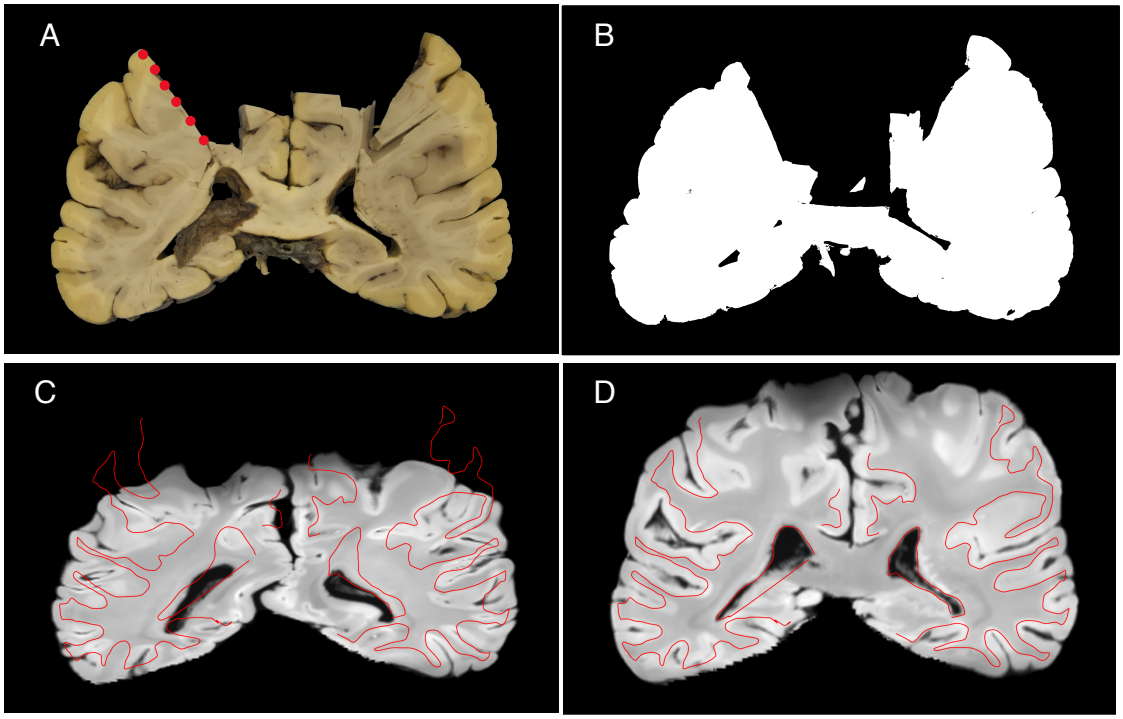

Figure 6.18: Stage-3 registration of coronal slices affected by M1S1 sampling. **(A)** Coronal slice photograph with tissue void due to M1S1 sampling. The artificial edges (*red dots*) are marked manually and correspond to the lateral surface of the hand M1S1 sampling site. **(B)** Slice mask used for stage-3 registration. **(C)** Erroneous stage-3 registration result in the absence of masking. **(D)** Accurate stage-3 registration result using the binary mask in (B). The *red contour* is an overlay of the gray-white matter boundary on the brain slice photograph.

Figure 6.18 shows a coronal slice that is severely damaged due to its intersection with the sampling sites of two M1S1 blocks (leg and hand) on each side. Using appropriate slice masking (excluding all artificial edges), such slices were registered to MRI space with the same accuracy as non-damaged slices using the stage-3 registration routine with the configurations that were described in 6.1.2. One of the artificial edges of the slice that corresponded to the lateral surface of the hand M1S1 block was manually marked with a few points (see Figure 6.18). This process was repeated on the next coronal slice that also intersected with the same M1S1 sampling site. These points were then mapped into MRI space by the optimised stage-3 transformation chains to indicate the position of the lateral surface of the M1S1 sampling site. Taking the mean and the 3rd principal component of the point cloud coordinates, the approximate centre and the normal vector of the lateral block surface could be determined. To initialise the registration, the centre of the tissue block (i.e. the source of the histology) was assumed to be 5 mm away from the lateral surface in the direction of medial-pointing normal vector. Figure 6.19 shows 3-D renderings of this initialisation process in MRI space.

The initialisation parameters (position and normal vector) were added to the stage-3 configurations, the tissue block photograph was converted to grayscale using the YIQ method, and both input images were normalised by the largest intensity value. A special binary mask was created for the 2-D input by thresholding the image at 10% to exclude most of the background area around artificial edges. The background around the pial surface was not masked, because this is a very important feature that proved to be necessary for a correct alignment. As a difference from the default stage-3 protocol, no mask was used for the MR volume.

The grid search in the first substage of stage 3 tested altogether 539 initialisations at 0.25 mm/px and the best 161 (30%) initialisations at native resolution (50 µm/px). The parameter grid comprised 11 equidistant positions along the normal vector in a 1 cm wide slab (i.e. every 1 mm) and 7×7 (internal) rotations about y and x in a range of $\pm 45°$ and $\pm 15°$, respectively. The different rotation ranges reflect an assumption that the rotation is the most variable about the sagittal axis (y).

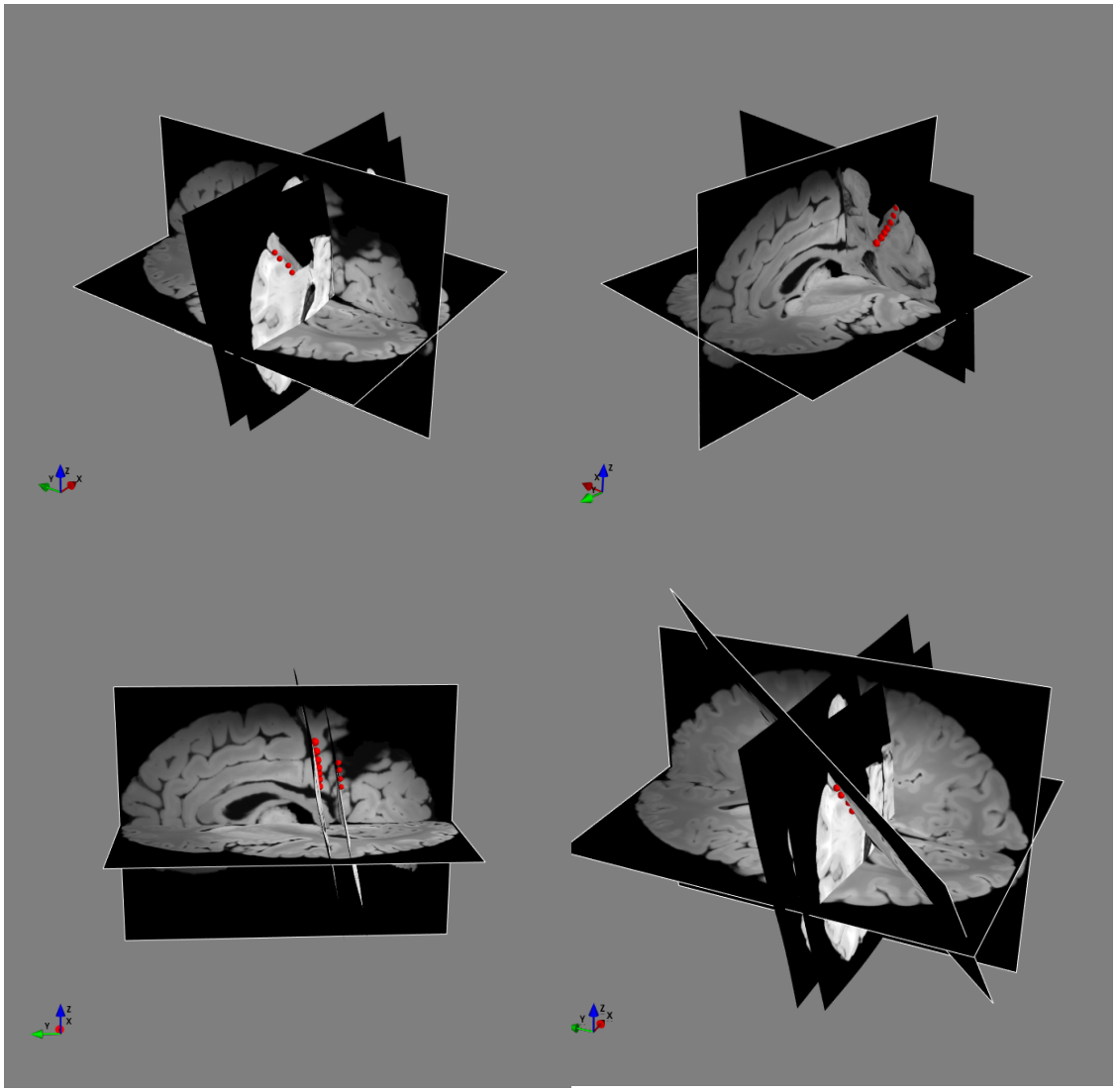

Figure 6.19: Different 3-D views on the manual tissue block initialisation process. The red dots were annotated in the pixel space of two neighbouring coronal slice photographs, and transformed to MRI space together with the slices. The bottom-right image shows a plane that was fitted to the transformed points to illustrate the the non-standard orientation of M1S1 block sampling.

This can be justified on the basis that M1S1 blocks are sampled such that they are (roughly) orthogonal to both the surface of the hemisphere and the central sulcus. Consequently, their orientation about the sagittal axis may vary in the $(-90°, 90°)$ range, depending on which part of the cortex they come from between the face area of the left hemisphere and the face area of the right hemisphere. The z rotations were optimised as part of the 2-D optimisation step after each initialisation from 0° in a range of ±20°, and the $z - y - x$ rotations were refined in the 3-D

optimisation step in the following range: $\pm\begin{bmatrix}5° & 10° & 10°\end{bmatrix}$. Scaling was fixed, and the 3-D translation parameters were optimised in a range of ±2.5 mm. In all cases the BOBYQA optimiser was used to minimise the MIND cost (the descriptors were calculated after resampling to 2-D, based on the 4 immediate neighbours). The affine substage was run at native resolution, using default parameter bounds. The non-linear substages optimised all xy and xyz displacements at native resolution, using 8 control points that were generated with the *Halton* scheme. The bounds on the displacements were set to ±1 mm for each component. Neither of the non-linear substages employed additional regularisation. The result of the fourth substage was accepted as the output of the pipeline.

The stage-1 and stage-3 transformation chains were combined to create a pixel-to-voxel mapping from the original histology image to the MR volume. The combined transformation chain is equivalent to that of the 3-stage pipeline, but the stage-2 chain was omitted. The MR data was finally resampled at 8 µm/px resolution onto the transformed histology domain.

6.2.3 Result

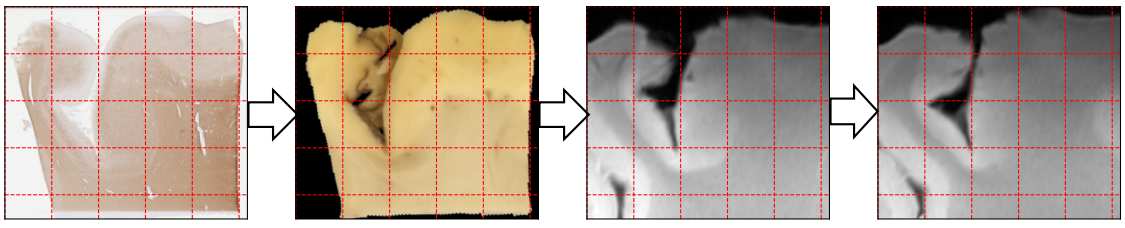

Figure 6.20: M1S1 registration result. The tissue section was taken from the hand area of the left motor cortex and immunostained for PLP. The 2-D MR images were resampled at the 8 µm/px resolution of the histology image from the structural (TRUFI) volume. Post-hoc optimisation (stage 4) was applied to minimise the anatomical discrepancy in the postcentral gyrus (left). Grid spacing: 5 mm.

Figure 6.20 shows the result of applying the M1S1 registration pipeline to a PLP-stained histological section from a left M1S1 block (hand area). The combined result after stage 3 seemed to be affected by an anatomical inconsistency between the histology and the tissue block images in the postcentral gyrus. This was

successfully compensated for by running the post-hoc optimisation (stage 4) with the previously described parameters (see Section 6.1.3).

6.2.4 Discussion

The difficulty of accurately aligning a small histology section with limited anatomical features can be better understood by looking at a random selection of the intermediate images that the post-hoc optimisation generated (Figure 6.21).

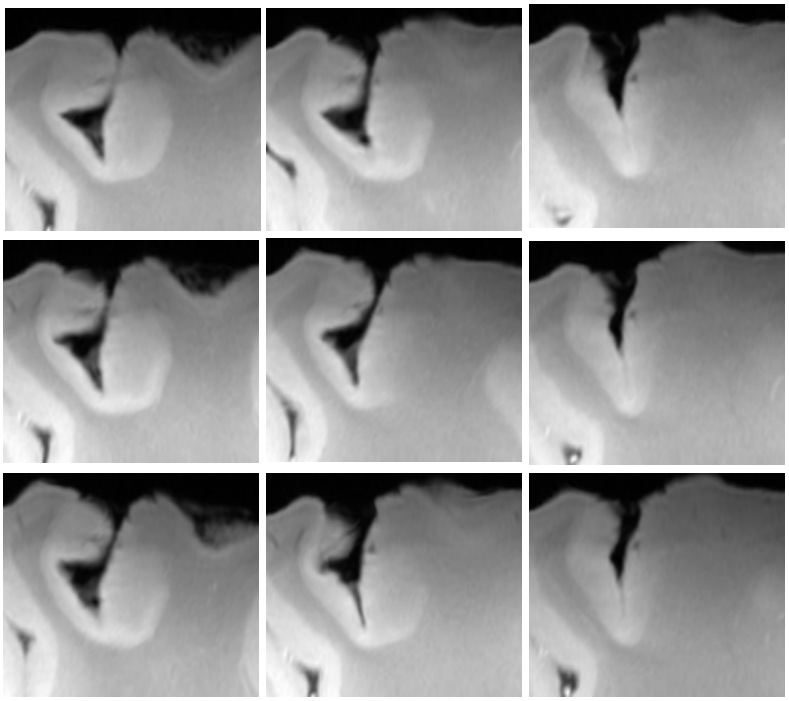

Figure 6.21: Small perturbations of the registered image during post-hoc optimisation (stage 4) lead to large changes in the appearance of the region, illustrating the difficulty of the registration.

All of these images correspond to less than ±1 mm translational and/or ±5° rotational offset, but the appearance of the M1S1 cross section is strikingly different even within these narrow bounds. This observation justifies the use of many initialisations during the stage-3 registration, and also why non-linear components are suppressed during the post-hoc optimisation. The fact that the block was initialised on the basis of stage-3 registrations, and the subsequent application of the M1S1 pipeline provided a visually accurate registration all provide evidence for the versatility, accuracy and robustness of the TIRL registration routines.

To further emphasise the need for accurate histology-to-MRI registration, we related the result of the M1S1 registration to a manually defined 3-D ROI. This was a registered version of the Brodmann Area 4 label of the FSL-Juelich Histological Atlas [101], that was trimmed to the hand area and adapted to the specific anatomy of the subject to analyse the correlation between MR and histology data in the motor cortex in one of our previous studies. Figure 6.22 shows the spatial relationship of the registered histology slice and the ROI.

The registered slice and the ROI do have an intersection, which indicates that the hand area was correctly identified by both. On a closer inspection however, one can see that the ROI is restricted to the white matter and also involves regions that are relatively distant from the slice, protruding on both sides, and towards deeper white matter regions.

In fact, the ROI had almost no intersection with the motor cortex, where the signs of iron overload or neuron loss would be expected [24]. If a comparison were erroneously made between the histology data of the gray matter and the MR properties of the white matter, correlations would probably look inconclusive, where in fact the absence of correlation may well be an important early disease sign, indicating the selective involvement [102], [103] of one of these compartments. The lateral extent of the ROI is also crucial, as the involvement of the motor cortex in ALS is not uniform [104].

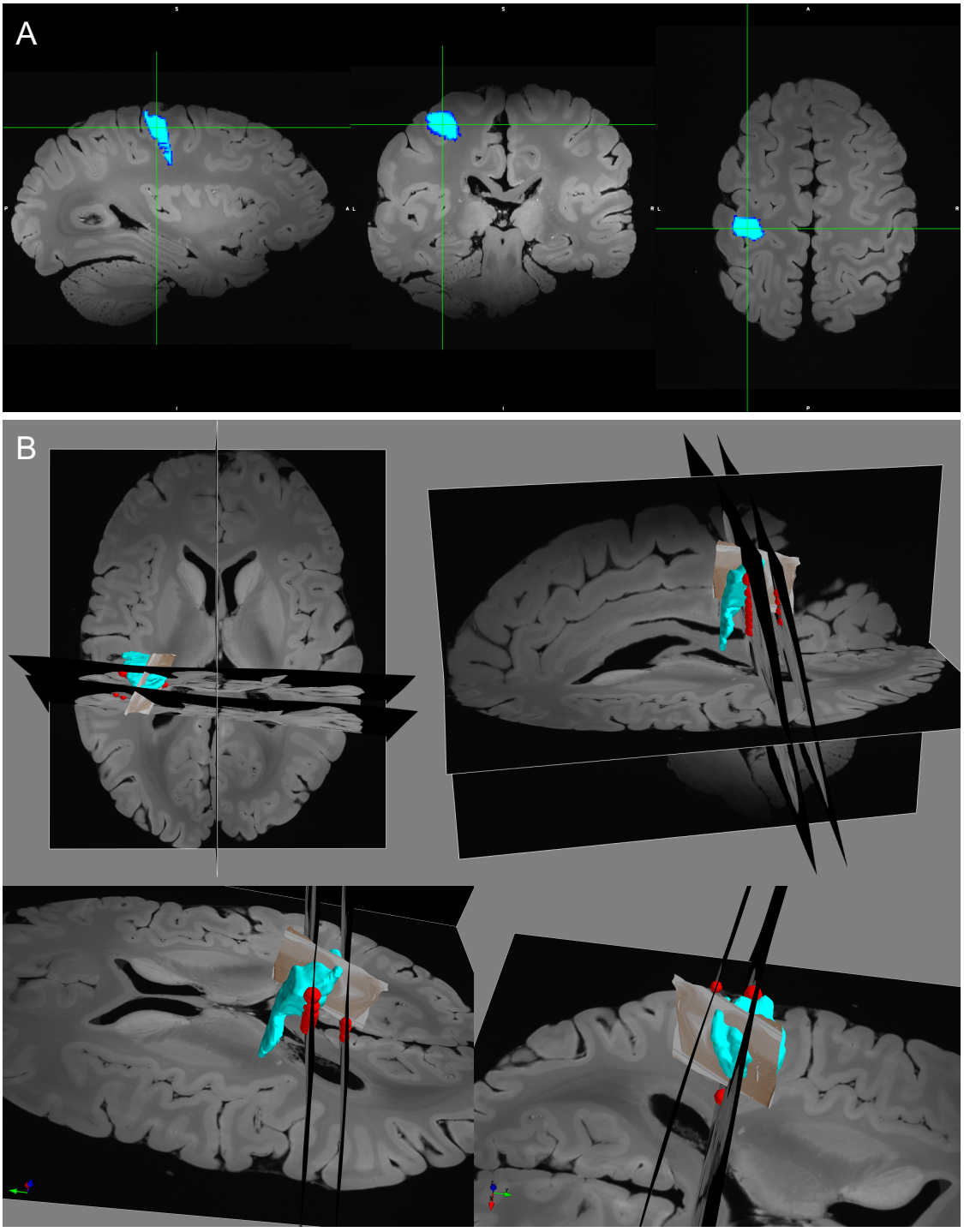

Figure 6.22: A 3-D ROI was created from the Brodmann Area 4 label from the Juelich Atlas. **(A)** 3-D manually defined ROI visualised in FSLeyes. **(B)** Alignment between the registered histology image and the manually-created 3-D ROI. The alignment is close, but there is no overlap in the region of the motor cortex, only the white matter. The extent of the ROI is much larger than the histology section.

7 Concluding remarks

7.1 Thesis contributions

All research that was presented in this thesis was motivated by finding a solution to a real-life problem, which is to register sparsely sampled histology and post-mortem MRI data in an existing dataset. This work draws on existing image registration concepts, but also makes important contributions to the two broader fields that it belongs to: medical image registration and neuroimaging in ALS. The main novel contributions are summarised below.

1. The thesis provides the first comprehensive evaluation of the MIND similarity metric in the context of histology-to-MRI registration. The flexibility of a novel image registration framework (TIRL) is demonstrated by prototyping four registration routines (stages 1-4) working with a range of image types and formats. The transformation chain formalism of ANTs is extended in TIRL, allowing both absolute and relative definitions of non-linear transformations and full interoperability with FSL transformations.

2. The TIRL stage-1 routine reimplements Heinrich's MIND-based deformable registration model with newly added image pre-processing and linear pre-registration steps. It was shown that this combined algorithm achieves superior

accuracy in the registration of histology to rigid reference images versus ANTs SyN with mutual information. The method was shown to work successfully with several types of histology stains.

3. The TIRL stage-2 routine presents a robust, automated, part-to-whole matching and registration algorithm that is based on the pairwise registration of a series of photographs. It was demonstrated that MIND can more reliably discriminate between tissue block-face photographs from different anatomical regions than mutual information. It was demonstrated that the stage-2 algorithm can simultaneously identify and register as many as 8 tissue block-face photographs with a single whole brain slice photograph in a fully automated way.

4. The TIRL stage-3 routine implements a novel multi-step optimisation algorithm for slice-to-volume registration using scattered control points and the MIND similarity metric. A comprehensive analysis of control point placement indicated diminishing returns in terms of registration accuracy versus control point count above 32 points. The Halton scheme is used for the first time for control point placement and it is shown to outperform regular control point placement. Finally, it was demonstrated that the algorithm can robustly estimate several millimetres of through-plane deformations of a slice and achieve sub-millimetre accuracy.

5. Evidence was provided that freehand dissection may lead to large deflections from an assumed slicing plane, which is often overlooked. To the best of our knowledge, TIRL is the first available platform that has been tested to compensate for this effect.

6. Registration by the stage-3 routine was shown to be more accurate than manual slice matching by landmarks.

7. Together the registration routines provided with TIRL constitute a fully functional automated histology-to-MRI registration pipeline, that achieved

sub-millimetre accuracy in the alignment of stand-alone histology images to volumetric MR data. The proposed method does not require specialised cutting or stain automation hardware for tissue processing and reduces the imperfections of alignment that arise from freehand brain cutting, which altogether make it suitable for integration into routine neuropathology practice. To the best of our knowledge, TIRL is the only registration method to integrate automated part-to-whole matching and slice-to-volume registration of conventional histology images from sparse, isolated anatomical sites. It is therefore an important step towards large-scale histological validation of imaging studies in healthy or diseased conditions, such as vascular dementia, Alzheimer's disease, Parkinson's disease, or ALS.

7.2 Future work

7.2.1 Scientific and technical developments of TIRL

Computational performance

Generality and optimal computational performance are often competing demands when implementing algorithms. Several features have been incorporated in TIRL to make computations more effective, such as parallel processing, chunked interpolation, array caching, optimising sub-chains of linear transformations by affine replacement, and avoiding interpolation of displacement fields where the field is defined over the same domain as the image. That said, greater emphasis was put on preserving the generality of the framework. Therefore, some of the computations may benefit from further optimisation in the future.

One example would concern the MIND similarity metric, which needs to be evaluated after each parameter update. This computation becomes a significant bottleneck for slice-to-volume registration at 0.25 mm/vx and above. The current implementation of MIND calculates patchwise distances using Gaussian convolution via Fast Fourier Transform instead of direct pixel arithmetic, as suggested by Heinrich et al [50]. It is anticipated that the same implementation on a GPU could lead to at least 3-4x faster performance (accounted for moving data between the

CPU and the GPU), leading to fully deformable slice-to-volume registrations in approximately 20 minutes at 0.25 mm/vx resolution.

Segmentation and registration by neural networks

Segmentation by deep learning could be useful to improve the proposed histology-MRI registration pipeline. Training networks to recognise histology artefacts, tissue voids or any unmatched side surface in photographs, as well as MRI artefacts would be useful to automate the masking on problematic images, the registration of which currently relies on manual masking as part of the pre-processing.

Neural networks have the capacity to speed up previously long iterative procedures (e.g. the entire FreeSurfer pipeline [105]) by learning the internal structure of the input and output data, and learning the space of typical mappings between these two. Consequently, when the input is supplied, a typical mapping can be generated very quickly by the trained model. A future study should explore whether a neural network could achieve fast slice-to-volume registration. Even a considerably close initialisation by the neural network would be sufficient to speed up this stage. Either of these would be an important element for the wider adoption of the TIRL framework on clinical grounds. The result of the stage-3 registrations presented in the thesis could provide the necessary ground truth that is required for the training. As the contrast is similar between the MR slices and the photographs, it may be useful to generate a very large amount of training data using simulated slices. Alternatively, the stage-3 registration results could be used to learn the synthesis of photographic images from simulated slices, and perform the training on those.

7.2.2 Registration without photographic intermediates

In Chapter 6 examples were shown how the pipeline can be used to register histology and MRI using either one or two photo intermediates. Recently we found that the stage-3 registration routine can be used effectively without photographic intermediates, if the distortions of the histology images are minimal. As part of an ongoing internal collaboration at the University of Oxford, stand-alone histology

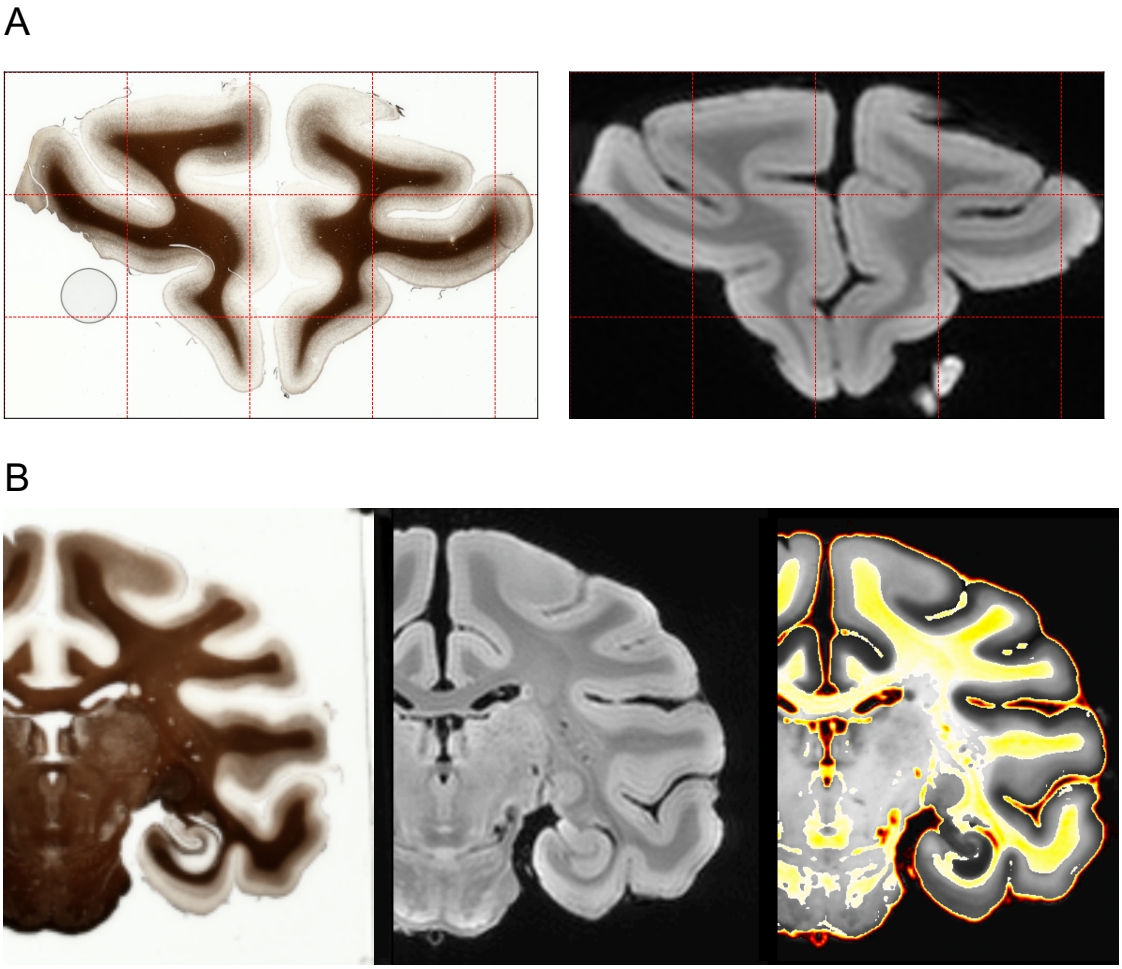

Figure 7.1: **(A)** Whole-brain histology section of the frontal lobe registered to MRI. The grid spacing is 10 mm. **(B)** Half-brain section close to the centre of the brain and the registered, resampled MR slice. The last image shows the histology image as an intensity overlay to visualise the precision of the alignment. Both histology images were prepared by the Gallyas silver impregnation technique to visualise myelinated fibres.

slices and polarised light micrographs were registered from the BigMac dataset [106] to a post-mortem gradient-echo image of a macaque brain at 0.3 mm/vx isotropic resolution. Representative examples can be seen in Figures 7.1 and 7.2. The macaque brain is small enough to be frozen and sectioned in a microtome to obtain whole-brain slices. As a result, the histology sections have minimal curvature and in-plane deformation with respect to the MR data, and stage-3 is executed with heavily restricted bounds on through-plane deformations.

This part of the BigMac dataset had been acquired without block-face photographs, hence TIRL's stage-3 routine was used as a means of direct microscopy-to-

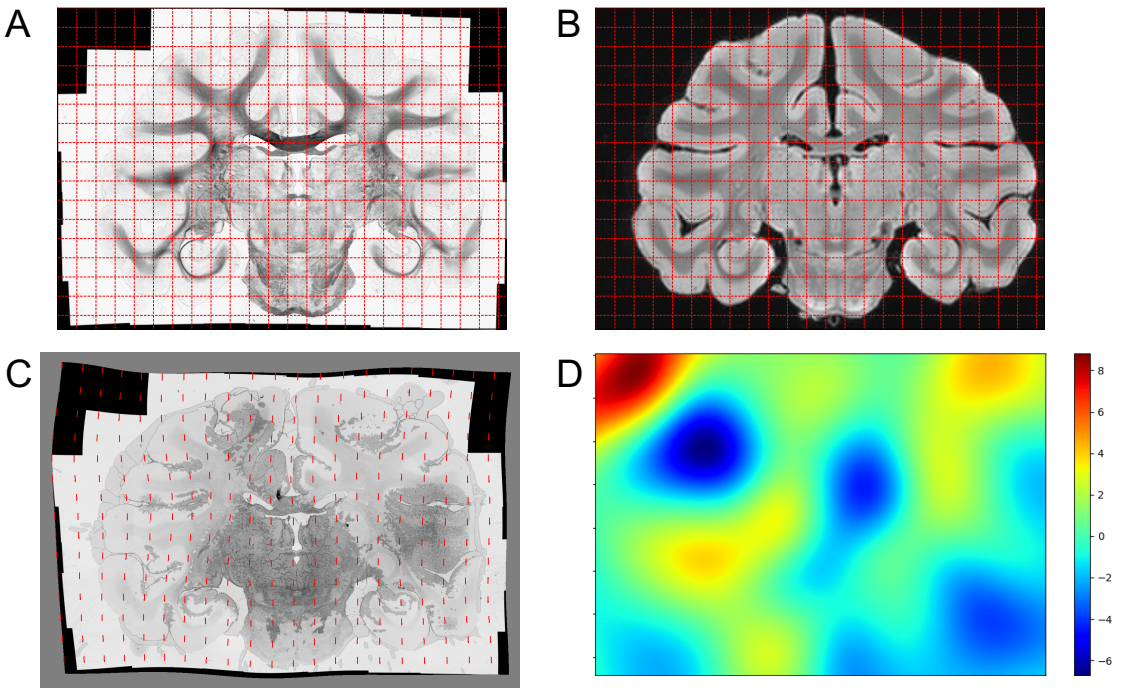

Figure 7.2: BigMac PLI. A: Transmittance image, B: Registered MRI slice after in-plane deformations. Grid spacing: 2.5 mm C: Inclination map with rotations that were introduced by TIRL: these need to be taken into account when comparing in-plane fibre orientations with diffusion maps. D: Local rotations shown in degrees.

MRI registration for preliminary analysis. In the future, more slices of the BigMac dataset will be acquired and registered with TIRL. These individual registrations will be used to obtain a robust local estimate of the slicing direction, and the MR volume itself will be used as the rigid reference for the reconstruction of the histology stack. This will require the alignment of histology slices with 0.3 mm spacing, avoiding any overlaps with neighbour slices. For simultaneous multi-slice registration, the stage-3 routine will be improved such that its cost function imposes constraints on the transformation parameters of consecutive slices.

7.2.3 Histology-informed biomarkers

Establishing novel imaging-derived biomarkers [107], [108] that can sensitively and specifically indicate the presence of a disease is one of the chief goals in modern medical imaging. Classic radiological signs such as signal hypo- and hyperintensities in weighted MRI scans have suboptimal disease specificity due to the complex dependency of the MRI signal on both the acquisition parameters and a spectrum

of elementary disease-related changes in tissue microstructure. By modelling the MR signal behaviour in the healthy and the diseased state of tissue, advanced microstructural MRI methods can be more specific to these elementary changes, and thus the underlying pathological process. The clinical translation of these methods requires thorough validation against histopathology, which will hopefully be facilitated by the availability of this MRI-histology registration tool.

An exciting future implication is that as soon as suitably large MRI/histology datasets become available, these could be parsed by neural networks to learn the patterns of subtle changes of the MRI signal that encode information about tissue pathology (as represented by various histology stains). Note that this concept is different from applying the same technique on existing hybrid datasets like BigBrain [76], because they are by no means representative of any disease process, or employ multi-modal histology. Instead of learning the generative model that synthesises MR data from a single histology modality or vice versa, the goal is to learn the representation of tissue-level pathophysiological processes in the altered MRI signal, such that when a characteristic pattern of multi-modal MR information is observed, it would be indicative of an underlying pathological process. This new generation of histology-inspired imaging biomarkers could be more sensitive and specific predictors of disease. In neurodegenerative conditions, increased sensitivity to the early sub-clinical stages of the disease is critical, as the anticipated benefit from any therapeutic approach is proportional to the remaining functional capacity of the central nervous system.

However, histology-inspired biomarkers cannot emerge from post-mortem data alone, as they need to be detected in clinical scans, and post-mortem images are not directly representative of the living state. Hence, we designed an in-vivo extension to the post-mortem ALS study (Figure 7.3), for which we have already started collecting data, but the study was put on hold due to infection control regulations in response to the COVID-19 pandemic.

The idea is to measure quantitative MRI properties in a short 30-minute session that is tolerable by people with symptomatic ALS. The session includes structural

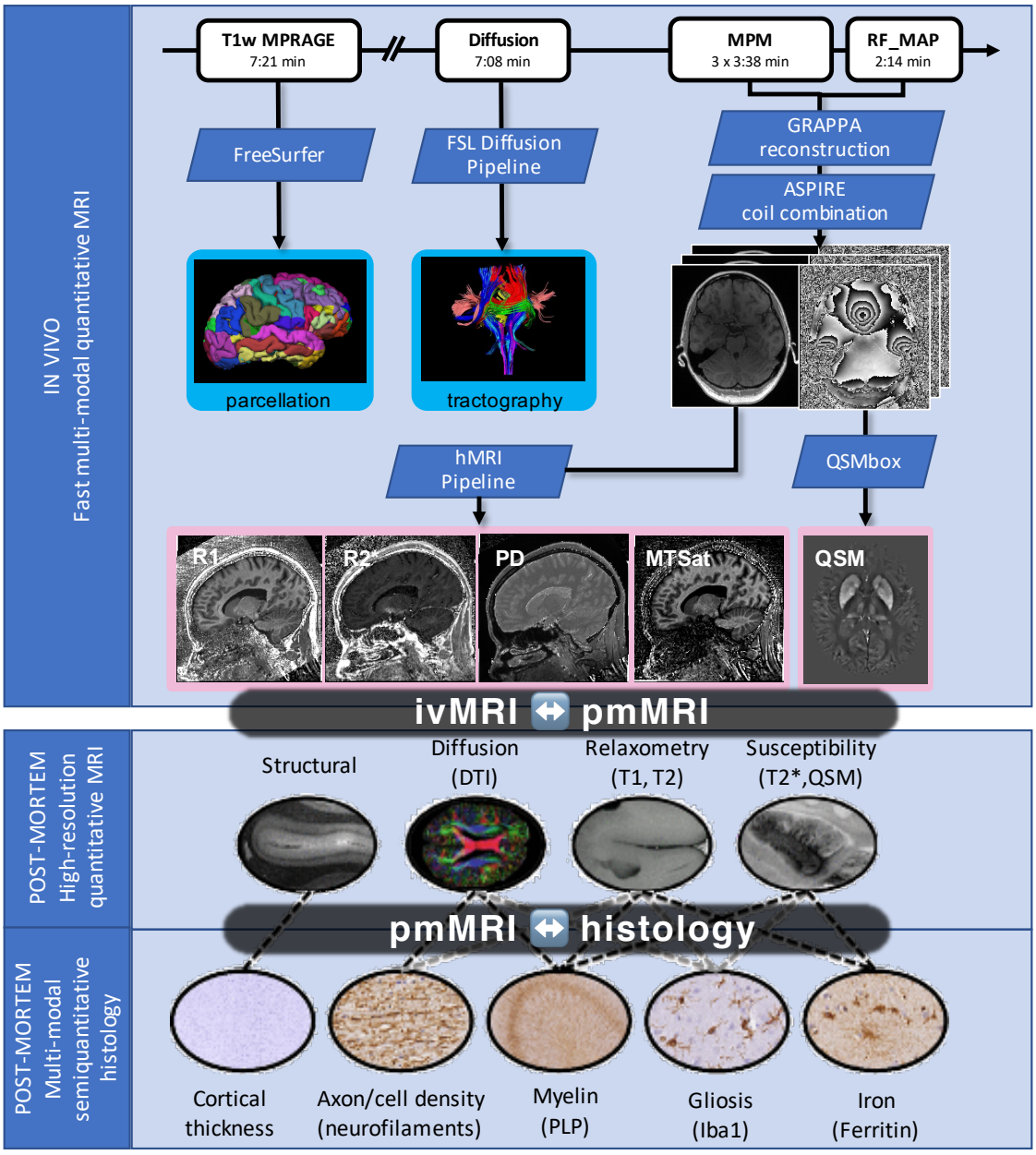

Figure 7.3: Integration of multi-modal quantitative in-vivo MRI (ivMRI), post-mortem MRI (pmMRI) and semi-quantitative histology data in a single subject (see details in the main text).

imaging (T1-weighted MPRAGE), a multi-shell flip-reversed diffusion protocol from BioBank [109], and a shortened version of the "in vivo histology" (hMRI) multi-parametric mapping protocol [110], [111] to derive diffusion-related parametric map (e.g. fractional anisotropy, fibre orientation distribution etc.) along with relaxometry (R1, R2*), proton density, magnetisation transfer saturation and quantitative susceptibility maps. Post-mortem follow-up imaging could be organised for consented individuals who wish to donate their brain for research after they reach the terminal stage of ALS. This way fully correspondent sets of in-vivo MRI (ivMRI), post-mortem MRI (pmMRI) and multi-modal histology data could be obtained that is representative to the evolution of a disease. This can be analysed for mutual correspondences between the modalities and eventually lead to histology-informed biomarkers.

7.3 Conclusion

The presented developments of the TIRL registration framework and the histology-to-MRI registration pipeline, together with the aforementioned future plans are anticipated to have an impact beyond their current application on the post-mortem ALS dataset. The tools are flexible enough to be applied in different research contexts, providing substantial grounds for future studies that aim to validate radiological findings against post-mortem histology data. Hopefully the software will also facilitate new explorative research into histology-informed microstructural imaging biomarkers that can be more sensitive predictors of neurodegeneration.

Appendices

A

Tensor Image Registration Library

A.1 What is TIRL?

Tensor Image Registration Library (TIRL) is a newly developed software platform[1] for creating bespoke image registration pipelines using the Python programming language. In the context of this thesis, TIRL is the computational and syntactic backend for all scripts that were written for the purpose of aligning tissue sections with volumetric MR data. In the broader sense, TIRL is a general-purpose Python library that conceptualises the image registration process as an orchestrated interplay between various components, which can be modified, replaced, and rearranged in a modular fashion to create and adapt an image registration pipeline to a particular problem. TIRL is a fully open-source project distributed as part of the FMRIB Software Library (FSL) v6.0.4, with the intention to be used and further extended in the future by both its developers and fellow researchers. The general structure of the library is illustrated in Figure A.1 and a description of its components is given below.

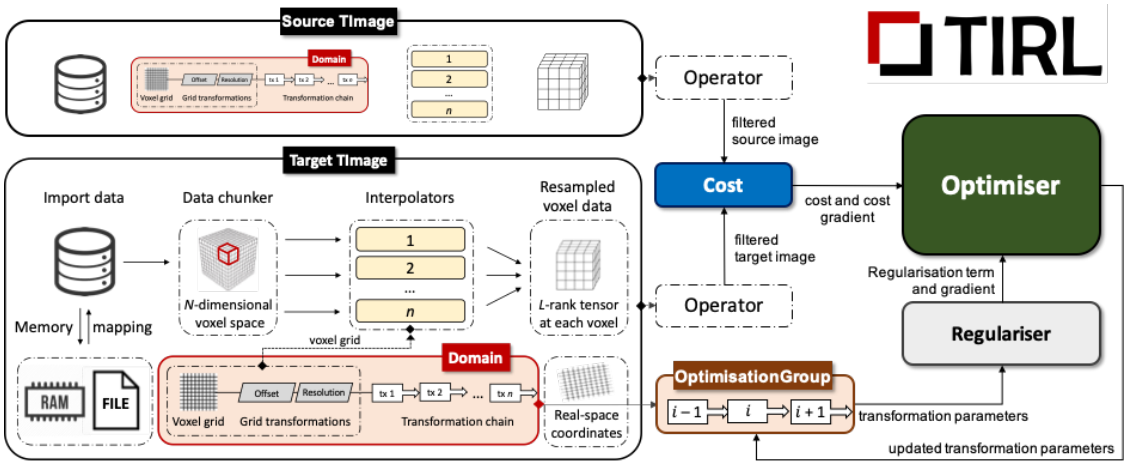

Figure A.1: Schematic illustration of the registration process in TIRL with the major library components involved (see main text for a description of each).

A.2 What are the components in TIRL?

TIRL follows an object-oriented programming paradigm. The registration process is realised by the interaction of several *objects*: the source and the target images (`TImage`), their domains (`Domain`), the coordinate transformations (`Transformation`), the cost function (`Cost`), the regularisation term (`Regulariser`), and the optimisation algorithm (`Optimiser`). TIRL defines custom file formats for saving any of the `TImage` (.timg), `Domain` (.dom), and `Transformation` (.tx) objects, which can later be loaded back into any compatible pipeline or used to extract quantitative transformation details after the registration is complete.

Similarly to a NIfTI image [29], a `TImage` is also defined in physical space. This means that the image data is stored in a regular voxel grid, but each voxel has an associate location in physical space (these are often referred to as "millimetre coordinates"). In the NIfTI standard, the mapping between voxel and mm coordinates is specified by an *affine matrix* (either the qform or the sform). TIRL generalises the mapping such that an arbitrary combination of linear and non-linear transformations can be assigned to the `Domain` of the `TImage`. The attached transformations constitute the *transformation chain* (`Chain`), that represents the mapping between voxel space and physical space.

[1]Python 3 library (approximately 30,000 lines of instructions) distributed via FSL. Author: Istvan N Huszar. Copyright: University of Oxford 2018-2020. Free for personal and academic use.

Any element of the transformation chain can be replaced or removed any time. The chains can also be split or concatenated, or used with different images. They consist of two parts: a set of *internal* transformations (gray parallelograms in figure A.1) and a set of *external* transformations (white rectangles). The internal part of the chain is entirely managed by TIRL, whereas the transformations in the external part need to be added and configured by the person who develops the registration routine. Despite this distinction, both internal and external transformations contribute equally to the mapping. Internal transformations are added automatically to the chain to preserve the position of the `TImage` in physical space when either the resolution of the `TImage` is changed (i.e. the image is downsampled to a sparser voxel grid or upsampled onto a denser grid), or the `TImage` is sliced (e.g. 100 pixels are removed from the top). With the added internal transformations the resolution of an image can be changed even after the registration is finished, and regions of interests can be extracted for further alysis without destroying the alignment of the image.

There is a variety of linear and non-linear transformations implemented in TIRL. Linear transformations include translation, rotation, shear, scale and affine. Rotations can be parametrised using any of the following: Euler angles (axes can be permuted), elements of a rotation matrix, a rotation axis and an angle, or quaternions. Conversion is implemented between any pair of these representations. The available options for non-linear transformation are polynomial fields, dense displacement fields and displacement fields with scattered control points and radial basis functions. Further options can be implemented by subclassing the `Transformation` base class.

With respect to similarity metrics, MSD, NMI and MIND are available, and further ones can be implemented by subclassing the `Cost` base class.

`Regularisers` are very similar to the `Cost` objects, as in they both compute a scalar value (and derivatives with respect to the transformation parameters if the optimisation is gradient-based), but the input of the `Regulariser` is the set of transformation parameters that are being optimised, not the images that are being registered. Available options for the `Regulariser` are the L1- and L2-norms

of the transformation parameters, diffusion regularisation, which penalises the first spatial derivatives of the transformations, and bending energy, where the penalty is on the second derivatives.

The `Optimiser` class has various subtypes for gradient-based and gradient-free optimisation; these are provided by the algorithms of the SciPy [112] and the NLOpt [113] Python libraries, or by custom implementation (e.g. the Gauss-Newton method). The `Optimiser` takes two arguments: the `Transformation`, the parameters of which will be updated, and a cost function, which is defined as any combination of `Cost` and `Regulariser` objects. Hybrid cost functions that are both landmark-based and intensity-based can be created this way, albeit at the expense of higher computational cost.

The registration commences by executing the "run" method of the `Optimiser`. At every iteration, the optimiser updates the transformation parameters, and issues the evaluation of the cost function, through the `Cost` and `Regulariser` objects. The cost and regularisation terms are therefore calculated in parallel. For the calculation of the cost terms, it is necessary for the images to have identical voxel shapes. This is achieved by resampling the source `TImage` on the `Domain` of the target `TImage`: the target `Domain` calculates the coordinates where the source image should be evaluated by its own `Interpolator`s. The `Interpolator`s are customisable, and for large enough images, they are even parallelised. It is the interpolated data and the unchanged target image data that are used by the `Cost` object to determine the goodness of the alignment between the images. `Operator`s are optional parts of the cost evaluation: they can be specified to pre-filter the images, or to calculate the MIND descriptors before a voxelwise intensity difference is taken by the `Cost`. The stopping criteria are encoded in the `Optimiser`, and if any of them is met, the registration concludes.

A.3 What makes TIRL a unique framework?

What makes TIRL unique is that any subset of transformations can be optimised simultaneously, if the the desired `Transformation` objects are encapsulated in an

`OptimisationGroup`. The `OptimisationGroup` mimics the API of the `Transformation` class, hence it is compatible with all `Optimisers`.

Another unique feature of TIRL is memory mapping. By default, image data is stored in the system memory. If it is too large, TIRL detects this and switches into memory mapping mode, which means that the image data will be accessed from the disk on demand, but only the necessary chunk is loaded at once not to overload the memory.

Finally, the name of the library comes from the very general image concept of the `TImage` object: "L-rank tensors on an N-dimensional domain with metadata'. With such a general assumption, the `TImage` aims to be a generic data container for the various image formats that a registration pipeline occasionally needs to work with.

References

[1] J. S. Salameh, J. Brown R. H., and J. D. Berry, "Amyotrophic lateral sclerosis: Review," *Semin Neurol*, vol. 35, no. 4, pp. 469–76, 2015. DOI: 10.1055/s-0035-1558984.

[2] K. Talbot, "Motor neurone disease," *Postgrad Med J*, vol. 78, no. 923, pp. 513–9, 2002. DOI: 10.1136/pmj.78.923.513.

[3] "2020 alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 16, no. 3, pp. 391–460, 2020. DOI: 10.1002/alz.12068.

[4] T. Pringsheim, N. Jette, A. Frolkis, and T. D. Steeves, "The prevalence of parkinson's disease: A systematic review and meta-analysis," *Movement Disorders*, vol. 29, no. 13, pp. 1583–1590, 2014. DOI: 10.1002/mds.25945.

[5] J.-M. Charcot and A. Joffroy, *Deux cas d'atrophie musculaire progressive : avec lÈsions de la substance grise et des faisceaux antÈrolatÈraux de la moelle ÈpiniËre*. Paris: Masson, 1869.

[6] C. G. Goetz, "Amyotrophic lateral sclerosis: Early contributions of jean-martin charcot," *Muscle Nerve*, vol. 23, no. 3, pp. 336–43, 2000. DOI: 10.1002/(sici)1097-4598(200003)23:3<336::aid-mus4>3.0.co;2-l.

[7] J. Brettschneider, K. Del Tredici, J. B. Toledo, J. L. Robinson, D. J. Irwin, M. Grossman, E. Suh, V. M. Van Deerlin, E. M. Wood, Y. Baek, L. Kwong, E. B. Lee, L. Elman, L. McCluskey, L. Fang, S. Feldengut, A. C. Ludolph, V. M.-Y. Lee, H. Braak, and J. Q. Trojanowski, "Stages of ptdp-43 pathology in amyotrophic lateral sclerosis," *Annals of Neurology*, vol. 74, no. 1, pp. 20–38, 2013. DOI: 10.1002/ana.23937.

[8] H. Braak, J. Brettschneider, A. C. Ludolph, V. M. Lee, J. Q. Trojanowski, and K. Del Tredici, "Amyotrophic lateral sclerosis–a model of corticofugal axonal spread," *Nat Rev Neurol*, vol. 9, no. 12, pp. 708–14, 2013. DOI: 10.1038/nrneurol.2013.221.

[9] F. Geser, V. M. Lee, and J. Q. Trojanowski, "Amyotrophic lateral sclerosis and frontotemporal lobar degeneration: A spectrum of tdp-43 proteinopathies," *Neuropathology*, vol. 30, no. 2, pp. 103–12, 2010. DOI: 10.1111/j.1440-1789.2009.01091.x.

[10] R. Steinbach, N. Gaur, B. Stubendorff, O. W. Witte, and J. Grosskreutz, "Developing a neuroimaging biomarker for amyotrophic lateral sclerosis: Multi-center data sharing and the road to a 'global cohort'," *Front Neurol*, vol. 9, p. 1055, 2018. DOI: 10.3389/fneur.2018.01055.

[11] P. Bede and O. Hardiman, "Lessons of ALS imaging: Pitfalls and future directions — a critical review," *NeuroImage: Clinical*, vol. 4, pp. 436–443, 2014. DOI: 10.1016/j.nicl.2014.02.011.

[12] M. Mazón, J. F. Vázquez Costa, A. Ten-Esteve, and L. Martí-Bonmatí, "Imaging biomarkers for the diagnosis and prognosis of neurodegenerative diseases. the example of amyotrophic lateral sclerosis," *Front Neurosci*, vol. 12, p. 784, 2018. DOI: 10.3389/fnins.2018.00784.

[13] M. R. Turner, M. C. Kiernan, P. N. Leigh, and K. Talbot, "Biomarkers in amyotrophic lateral sclerosis," *Lancet Neurol*, vol. 8, no. 1, pp. 94–109, 2009. DOI: 10.1016/s1474-4422(08)70293-x.

[14] A. Chase, "Improved imaging biomarkers in amyotrophic lateral sclerosis," *Nature Reviews Neurology*, vol. 10, no. 8, pp. 427–427, 2014. DOI: 10.1038/nrneurol.2014.125.

[15] V. Grollemund, P. F. Pradat, G. Querin, F. Delbot, G. Le Chat, J. F. Pradat-Peyre, and P. Bede, "Machine learning in amyotrophic lateral sclerosis: Achievements, pitfalls, and future directions," *Front Neurosci*, vol. 13, p. 135, 2019. DOI: 10.3389/fnins.2019.00135.

[16] P. Bede, "The histological correlates of imaging metrics: Postmortem validation of in vivo findings," *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, vol. 20, no. 7-8, pp. 457–460, 2019. DOI: 10.1080/21678421.2019.1639195.

[17] H. A. Alturkistani, F. M. Tashkandi, and Z. M. Mohammedsaleh, "Histological stains: A literature review and case study," *Glob J Health Sci*, vol. 8, no. 3, pp. 72–9, 2015. DOI: 10.5539/gjhs.v8n3p72.

[18] K. Bera, K. A. Schalper, D. L. Rimm, V. Velcheti, and A. Madabhushi, "Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology," *Nature Reviews Clinical Oncology*, vol. 16, no. 11, pp. 703–715, 2019. DOI: 10.1038/s41571-019-0252-y.

[19] C. Wang, S. Foxley, O. Ansorge, S. Bangerter-Christensen, M. Chiew, A. Leonte, R. A. L. Menke, J. Mollink, M. Pallebage-Gamarallage, M. R. Turner, K. L. Miller, and B. C. Tendler, "Methods for quantitative susceptibility and r2* mapping in whole post-mortem brains at 7t applied to amyotrophic lateral sclerosis," *NeuroImage*, vol. 222, p. 117 216, 2020. DOI: 10.1016/j.neuroimage.2020.117216.

[20] S. Foxley, S. Jbabdi, S. Clare, W. Lam, O. Ansorge, G. Douaud, and K. Miller, "Improving diffusion-weighted imaging of post-mortem human brains: Ssfp at 7t," *NeuroImage*, vol. 102, pp. 579–589, 2014. DOI: 10.1016/j.neuroimage.2014.08.014.

[21] M. Pallebage-Gamarallage, S. Foxley, R. A. L. Menke, I. N. Huszar, M. Jenkinson, B. C. Tendler, C. Wang, S. Jbabdi, M. R. Turner, K. L. Miller, and O. Ansorge, "Dissecting the pathobiology of altered mri signal in amyotrophic lateral sclerosis: A post mortem whole brain sampling strategy for the integration of ultra-high-field mri and quantitative neuropathology," *BMC Neurosci*, vol. 19, no. 1, p. 11, 2018. DOI: 10.1186/s12868-018-0416-1.

[22] M. D. Meadowcroft, N. J. Mutic, D. C. Bigler, J. L. Wang, Z. Simmons, J. R. Connor, and Q. X. Yang, "Histological-mri correlation in the primary motor cortex of patients with amyotrophic lateral sclerosis," *J Magn Reson Imaging*, vol. 41, no. 3, pp. 665–75, 2015. DOI: 10.1002/jmri.24582.

[23] C. Langkammer, F. Schweser, N. Krebs, A. Deistung, W. Goessler, E. Scheurer, K. Sommer, G. Reishofer, K. Yen, F. Fazekas, S. Ropele, and J. R. Reichenbach, "Quantitative susceptibility mapping (qsm) as a means to measure brain iron? a post mortem validation study," *NeuroImage*, vol. 62, no. 3, pp. 1593–1599, 2012. DOI: 10.1016/j.neuroimage.2012.05.049.

[24] J. Y. Kwan, S. Y. Jeong, P. Van Gelderen, H. X. Deng, M. M. Quezado, L. E. Danielian, J. A. Butman, L. Chen, E. Bayat, J. Russell, T. Siddique, J. H. Duyn, T. A. Rouault, and M. K. Floeter, "Iron accumulation in deep cortical layers accounts for mri signal abnormalities in als: Correlating 7 tesla mri and pathology," *PLoS One*, vol. 7, no. 4, e35241, 2012. DOI: 10.1371/journal.pone.0035241.

[25] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003. DOI: 10.1016/S0262-8856(03)00137-9.

[26] D. Rueckert and J. A. Schnabel, "Medical image registration," in *Biomedical Image Processing*, T. M. Deserno, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 131–154. DOI: 10.1007/978-3-642-15816-2_5.

[27] J. V. Hajnal, D. J. Hawkes, and D. L. G. Hill, "Medical image registration," 2001.

[28] A. W. Toga and P. M. Thompson, "The role of image registration in brain mapping," *Image Vis Comput*, vol. 19, no. 1-2, pp. 3–24, 2001. DOI: 10.1016/s0262-8856(00)00055-x.

[29] *Neuroimaging informatics technology initiative*, Web Page. [Online]. Available: https://nifti.nimh.nih.gov/.

[30] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "Fsl," *Neuroimage*, vol. 62, no. 2, pp. 782–90, 2012. DOI: 10.1016/j.neuroimage.2011.09.015.

[31] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 6, pp. 567–585, 1989. DOI: 10.1109/34.24792.

[32] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast mr images," *IEEE Trans Med Imaging*, vol. 18, no. 8, pp. 712–21, 1999. DOI: 10.1109/42.796284.

[33] J. Ashburner and K. J. Friston, "Nonlinear spatial normalization using basis functions," *Hum Brain Mapp*, vol. 7, no. 4, pp. 254–66, 1999. DOI: 10.1002/(SICI)1097-0193(1999)7:4<254::AID-HBM4>3.0.CO;2-G.

[34] Y. Amit, "A nonlinear variational problem for image matching," *SIAM J. Sci. Comput.*, vol. 15, pp. 207–224, 1994.

[35] T. Kim, M. Singh, W. Sungkarat, C. Zarow, and H. Chui, "Automatic registration of postmortem brain slices to mri reference volume," *IEEE Transactions on Nuclear Science*, vol. 47, no. 4, pp. 1607–1613, 2000. DOI: 10.1109/23.873023.

[36] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *International Journal of Computer Vision*, vol. 61, no. 2, pp. 139–157, 2005. DOI: 10.1023/B:VISI.0000043755.93987.aa.

[37] B. B. Avants, N. J. Tustison, M. Stauffer, G. Song, B. Wu, and J. C. Gee, "The Insight ToolKit image registration framework," *Front Neuroinform*, vol. 8, p. 44, 2014. DOI: 10.3389/fninf.2014.00044.

[38] *Diffeomorphism*, Web Page. [Online]. Available: http://encyclopediaofmath.org/index.php?title=Diffeomorphism&oldid=46650.

[39] E. Ferrante and N. Paragios, "Slice-to-volume medical image registration: A survey," *Medical Image Analysis*, vol. 39, pp. 101–123, 2017. DOI: 10.1016/j.media.2017.04.010.

[40] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 2, 1150–1157 vol.2. DOI: 10.1109/ICCV.1999.790410.

[41] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey Vision Conference*.

[42] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, 886–893 vol. 1. DOI: 10.1109/CVPR.2005.177.

[43] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Computer Vision – ECCV 2006*, A. Leonardis, H. Bischof, and A. Pinz, Eds., Springer Berlin Heidelberg, pp. 404–417.

[44] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992. DOI: 10.1109/34.121791.

[45] T. Ohnishi, Y. Nakamura, T. Tanaka, T. Tanaka, N. Hashimoto, H. Haneishi, T. T. Batchelor, E. R. Gerstner, J. W. Taylor, M. Snuderl, and Y. Yagi, "Deformable image registration between pathological images and mr image via an optical macro image," *Pathol Res Pract*, vol. 212, no. 10, pp. 927–936, 2016. DOI: 10.1016/j.prp.2016.07.018.

[46] D. N. Greve and B. Fischl, "Accurate and robust brain image alignment using boundary-based registration," *Neuroimage*, vol. 48, no. 1, pp. 63–72, 2009. DOI: 10.1016/j.neuroimage.2009.06.060.

[47] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, J. Maurer C. R., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. Hill, D. J. Hawkes, C. Studholme, J. B. Maintz, M. A. Viergever, G. Malandain, R. P. Woods, and et al., "Comparison and evaluation of retrospective intermodality brain image registration techniques," *J Comput Assist Tomogr*, vol. 21, no. 4, pp. 554–66, 1997. DOI: 10.1097/00004728-199707000-00007.

[48] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Med Image Anal*, vol. 12, no. 1, pp. 26–41, 2008. DOI: 10.1016/j.media.2007.06.004.

[49] C. E. Shannon and W. Weaver, *The mathematical theory of communication*. Urbana: University of Illinois Press, 1963.

[50] 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," *Med Image Anal*, vol. 16, no. 7, pp. 1423–35, 2012. DOI: 10.1016/j.media.2012.05.008.

[51] G. Haskins, U. Kruger, and P. Yan, "Deep learning in medical image registration: A survey," *Machine Vision and Applications*, vol. 31, no. 1, p. 8, 2020. DOI: 10.1007/s00138-020-01060-x.

[52] K. Madsen, H. Nielsen, O. Tingleff, and D. t. u. I. o. M. Modellering, *Methods for Non-linear Least Squares Problems*. Informatics and Mathematical Modelling, Technical University of Denmark, 2004.

[53] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1153–1190, 2013. DOI: 10.1109/TMI.2013.2265603.

[54] M. J. D. Powell, "The newuoa software for unconstrained optimization without derivatives," in *Large-Scale Nonlinear Optimization*, G. Di Pillo and M. Roma, Eds. Boston, MA: Springer US, 2006, pp. 255–297. DOI: 10.1007/0-387-30065-1_16.

[55] M. Powell, "The bobyqa algorithm for bound constrained optimization without derivatives," *Technical Report, Department of Applied Mathematics and Theoretical Physics*, 2009.

[56] S. Osechinskiy and F. Kruggel, "Slice-to-volume nonrigid registration of histological sections to mr images of the human brain," *Anat Res Int*, vol. 2011, p. 287 860, 2011. DOI: 10.1155/2011/287860.

[57] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios, "Dense image registration through mrfs and efficient linear programming," *Medical Image Analysis*, vol. 12, no. 6, pp. 731–741, 2008. DOI: 10.1016/j.media.2008.03.006. [Online]. Available: https://dx.doi.org/10.1016/j.media.2008.03.006.

[58] D. Zikic, B. Glocker, O. Kutter, M. Groher, N. Komodakis, A. Kamen, N. Paragios, and N. Navab, "Linear intensity-based image registration by markov random fields and discrete optimization," *Medical image analysis*, vol. 14, pp. 550–62, 2010. DOI: 10.1016/j.media.2010.04.003.

[59] A. W. Toga, K. L. Ambach, and S. Schluender, "High-resolution anatomy from in situ human brain," *Neuroimage*, vol. 1, no. 4, pp. 334–44, 1994. DOI: 10.1006/nimg.1994.1018.

[60] T. Schormann, A. Dabringhaus, and K. Zilles, "Statistics of deformations in histology and application to improved alignment with mri," *IEEE transactions on medical imaging*, vol. 14 1, pp. 25–35, 1995.

[61] M. S. Mega, S. S. Chen, P. M. Thompson, R. P. Woods, T. J. Karaca, A. Tiwari, H. V. Vinters, G. W. Small, and A. W. Toga, "Mapping histology to metabolism: Coregistration of stained whole-brain sections to premortem pet in alzheimer's disease," *NeuroImage*, vol. 5, no. 2, pp. 147–153, 1997. DOI: 10.1006/nimg.1996.0255.

[62] T. Schormann and K. Zilles, "Three-dimensional linear and nonlinear transformations: An integration of light microscopical and mri data," *Human Brain Mapping*, vol. 6, no. 5-6, pp. 339–347, 1998. DOI: 10.1002/(SICI)1097-0193(1998)6:5/6<339::AID-HBM3>3.0.CO;2-Q.

[63] M. A. Jacobs, J. P. Windham, H. Soltanian-Zadeh, D. J. Peck, and R. A. Knight, "Registration and warping of magnetic resonance images to histological sections," *Medical Physics*, vol. 26, no. 8, pp. 1568–1578, 1999. DOI: 10.1118/1.598671.

[64] S. Ourselin, E. Bardinet, D. Dormont, G. Malandain, A. Roche, N. Ayache, D. Tandé, K. Parain, and J. Yelnik, "Fusion of histological sections and mr images: Towards the construction of an atlas of the human basal ganglia," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2001*, W. J. Niessen and M. A. Viergever, Eds., Springer Berlin Heidelberg, pp. 743–751.

[65] É. Bardinet, S. Ourselin, D. Dormont, G. Malandain, D. Tandé, K. Parain, N. Ayache, and J. Yelnik, "Co-registration of histological, optical and mr data of the human brain," in *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002*, T. Dohi and R. Kikinis, Eds., Springer Berlin Heidelberg, pp. 548–555.

[66] A. W. Anderson, A. S. Choe, I. Stepniewska, and D. C. Colvin, "Comparison of brain white matter fiber orientation measurements based on diffusion tensor imaging and light microscopy," *Conf Proc IEEE Eng Med Biol Soc*, vol. 2006, pp. 2249–51, 2006. DOI: 10.1109/iembs.2006.259554.

[67] J. Dauguet, T. Delzescaux, F. Condé, J. F. Mangin, N. Ayache, P. Hantraye, and V. Frouin, "Three-dimensional reconstruction of stained histological slices and 3d non-linear registration with in-vivo mri for whole baboon brain," *J Neurosci Methods*, vol. 164, no. 1, pp. 191–204, 2007. DOI: 10.1016/j.jneumeth.2007.04.017.

[68] 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," *Magn Reson Imaging*, vol. 29, no. 5, pp. 683–92, 2011. DOI: 10.1016/j.mri.2011.02.022.

[69] S. Yang, Z. Yang, K. Fischer, K. Zhong, J. Stadler, F. Godenschweger, J. Steiner, H. J. Heinze, H. G. Bernstein, B. Bogerts, C. Mawrin, D. C. Reutens, O. Speck, and M. Walter, "Integration of ultra-high field mri and histology for connectome based research of brain disorders," *Front Neuroanat*, vol. 7, p. 31, 2013. DOI: 10.3389/fnana.2013.00031.

[70] J. Pichat, J. E. Iglesias, T. Yousry, S. Ourselin, and M. Modat, "A survey of methods for 3d histology reconstruction," *Med Image Anal*, vol. 46, pp. 73–105, 2018. DOI: 10.1016/j.media.2018.02.004.

[71] G. Rohde, A. Aldroubi, and B. Dawant, *Adaptive-bases algorithm for nonrigid image registration*, ser. Medical Imaging 2002. SPIE, 2002, vol. 4684.

[72] J. E. Iglesias, M. Modat, L. Peter, A. Stevens, R. Annunziata, T. Vercauteren, E. Lein, B. Fischl, and S. Ourselin, "Joint registration and synthesis using a probabilistic model for alignment of mri and histological sections," *Medical Image Analysis*, vol. 50, pp. 127–144, 2018. DOI: 10.1016/j.media.2018.09.002.

[73] J. Pichat, J. Eugenio Iglesias, S. Nousias, T. Yousry, S. Ourselin, and M. Modat, "Part-to-whole registration of histology and mri using shape elements," *arXiv e-prints*, arXiv:1708.08117, 2017.

[74] M. Mancini, A. Casamitjana, L. Peter, E. Robinson, S. Crampsie, D. L. Thomas, J. L. Holton, Z. Jaunmuktane, and J. E. Iglesias, "A multimodal computational pipeline for 3d histology of the human brain," *bioRxiv*, p. 2020.02.10.941948, 2020. DOI: 10.1101/2020.02.10.941948.

[75] G. Malandain, E. Bardinet, K. Nelissen, and W. Vanduffel, "Fusion of autoradiographs with an mr volume using 2-d and 3-d linear transformations," *NeuroImage*, vol. 23, pp. 111–27, 2004. DOI: 10.1016/j.neuroimage.2004.04.038.

[76] K. Amunts, C. Lepage, L. Borgeat, H. Mohlberg, T. Dickscheid, M.-É. Rousseau, S. Bludau, P.-L. Bazin, L. B. Lewis, A.-M. Oros-Peusquens, N. J. Shah, T. Lippert, K. Zilles, and A. C. Evans, "Bigbrain: An ultrahigh-resolution 3d human brain model," *Science*, vol. 340, no. 6139, p. 1472, 2013. DOI: 10.1126/science.1235381.

[77] M. Alegro, E. A. Junior, B. Loring, H. Heinsen, E. Alho, L. Zöllei, D. Ushizima, and L. T. Grinberg, "Multimodal whole brain registration: Mri and high resolution histology," *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 634–642, 2016.

[78] M. Alegro, E. J. L. Alho, M. da Graca Morais Martin, L. Teneholz Grinberg, H. Heinsen, R. de Deus Lopes, E. Amaro-Jr, and L. Zöllei, "Automating whole brain histology to mri registration: Implementation of a computational pipeline," *arXiv e-prints*, arXiv:1905.09339, 2019.

[79] H. Ni, C. Tan, Z. Feng, S. Chen, Z. Zhang, W. Li, Y. Guan, H. Gong, Q. Luo, and A. Li, "A robust image registration interface for large volume brain atlas," *Scientific Reports*, vol. 10, no. 1, p. 2139, 2020. DOI: 10.1038/s41598-020-59042-y.

[80] M. Singh, A. Rajagopalan, T. S. Kim, D. Hwang, H. Chui, X. L. Zhang, A. Y. Lee, and C. Zarow, "Co-registration of in-vivo human mri brain images to postmortem histological microscopic images," *Int J Imaging Syst Technol*, vol. 18, no. 5-6, pp. 325–335, 2008. DOI: 10.1002/ima.v18:5/6.

[81] C. R. Meyer, B. A. Moffat, K. K. Kuszpit, P. L. Bland, P. E. McKeever, T. D. Johnson, T. L. Chenevert, A. Rehemtulla, and B. D. Ross, "A methodology for registration of a histological slide and in vivo mri volume based on optimizing mutual information," *Mol Imaging*, vol. 5, no. 1, pp. 16–23, 2006.

[82] S. Osechinskiy and F. Kruggel, "Deformable registration of histological sections to brain mr images using a hybrid boundary-based slice-to-volume approach," *Annu Int Conf IEEE Eng Med Biol Soc*, vol. 2011, pp. 4876–9, 2011. DOI: 10.1109/iembs.2011.6091208.

[83] A. De Barros, G. Arribarat, J. Combis, P. Chaynes, and P. Péran, "Matching ex vivo mri with iron histology: Pearls and pitfalls," *Frontiers in Neuroanatomy*, vol. 13, no. 68, 2019. DOI: 10.3389/fnana.2019.00068.

[84] P. Majka and D. K. Wójcik, "Possum—a framework for three-dimensional reconstruction of brain images from serial sections," *Neuroinformatics*, vol. 14, no. 3, pp. 265–278, 2016. DOI: 10.1007/s12021-015-9286-1.

[85] D. Adler and P. Yuskevich, *Histolozee*, Computer Program. [Online]. Available: http://picsl.upenn.edu/software/histolozee/.

[86] B. C. Tendler, S. Foxley, M. Hernandez-Fernandez, M. Cottaar, C. Scott, O. Ansorge, K. L. Miller, and S. Jbabdi, "Use of multi-flip angle measurements to account for transmit inhomogeneity and non-gaussian diffusion in dw-ssfp," *NeuroImage*, vol. 220, p. 117 113, 2020. DOI: 10.1016/j.neuroimage.2020.117113.

[87] B. C. Tendler, S. Foxley, M. Cottaar, S. Jbabdi, and K. L. Miller, "Modeling an equivalent b-value in diffusion-weighted steady-state free precession," *Magnetic Resonance in Medicine*, vol. 84, no. 2, pp. 873–884, 2020. DOI: 10.1002/mrm.28169.

[88] J. Bastacky, T. L. Hayes, and R. P. Gelinas, "Quantitation of shrinkage during preparation for scanning electron microscopy: Human lung," *Scanning*, vol. 7, no. 3, pp. 134–140, 1985. DOI: 10.1002/sca.4950070306.

[89] *Darktable*, Web Page. [Online]. Available: https://www.darktable.org/.

[90] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: An open-source platform for biological-image analysis," *Nature Methods*, vol. 9, no. 7, pp. 676–682, 2012. DOI: 10.1038/nmeth.2019.

[91] M. Jenkinson and S. M. Smith, "Optimisation in robust linear registration of brain images," Unpublished Work, 2000.

[92] J. Modersitzki, *Numerical methods for image registration*. Oxford: Oxford Univ Press, 2009.

[93] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979. DOI: 10.1109/TSMC.1979.4310076.

[94] D. Demirović, "An implementation of the mean shift algorithm," *Image Processing On Line*, vol. 9, pp. 251–268, 2019.

[95] D. W. Shattuck and R. M. Leahy, "Brainsuite: An automated cortical surface identification tool," *Medical Image Analysis*, vol. 6, no. 2, pp. 129–142, 2002. DOI: 10.1016/S1361-8415(02)00054-3.

[96] P. A. Yushkevich, G. Yang, and G. Gerig, "Itk-snap: An interactive tool for semi-automatic segmentation of multi-modality biomedical images," *Annu Int Conf IEEE Eng Med Biol Soc*, vol. 2016, pp. 3342–3345, 2016. DOI: 10.1109/embc.2016.7591443.

[97] P. McCarthy, *FSLeyes*, Computer Program, 2020. DOI: http://doi.org/10.5281/zenodo.3937147.

[98] J. H. Halton, "Algorithm 247: Radical-inverse quasi-random point sequence," *Communications of the ACM*, vol. 7, no. 12, pp. 701–702, 1964.

[99] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–41, 2002. DOI: 10.1016/s1053-8119(02)91132-8.

[100] D. Alexander, C. Pierpaoli, P. Basser, and J. Gee, "Spatial transformations of diffusion tensor magnetic resonance images," *IEEE transactions on medical imaging*, vol. 20, pp. 1131–9, 2001. DOI: 10.1109/42.963816.

[101] *Fsl-juelich histological atlas*, Dataset. [Online]. Available: https://neuro.debian.net/pkgs/fsl-juelich-histological-atlas.html.

[102] H.-P. Müller, F. Agosta, M. Gorges, R. Kassubek, E. G. Spinelli, N. Riva, A. C. Ludolph, M. Filippi, and J. Kassubek, "Cortico-efferent tract involvement in primary lateral sclerosis and amyotrophic lateral sclerosis: A two-centre tract of interest-based dti analysis," *NeuroImage: Clinical*, vol. 20, pp. 1062–1069, 2018. DOI: 10.1016/j.nicl.2018.10.005.

[103] M. Fatima, R. Tan, G. M. Halliday, and J. J. Kril, "Spread of pathology in amyotrophic lateral sclerosis: Assessment of phosphorylated tdp-43 along axonal pathways," *Acta Neuropathologica Communications*, vol. 3, no. 1, p. 47, 2015. DOI: 10.1186/s40478-015-0226-y.

[104] J. Ravits, P. Paul, and C. Jorg, "Focality of upper and lower motor neuron degeneration at the clinical onset of als," *Neurology*, vol. 68, no. 19, pp. 1571–5, 2007. DOI: 10.1212/01.wnl.0000260965.20021.47.

[105] L. Henschel, S. Conjeti, S. Estrada, K. Diers, B. Fischl, and M. Reuter, "Fastsurfer - a fast and accurate deep learning based neuroimaging pipeline," *NeuroImage*, vol. 219, p. 117 012, 2020. DOI: 10.1016/j.neuroimage.2020.117012.

[106] A. Howard and K. Miller, "The bigmac dataset: Ultra-high angular resolution diffusion imaging and multi-contrast microscopy of a whole macaque brain," in *Proceedings of the International Society for Magnetic Resonance Imaging in Medicine*, 2019.

[107] M. R. Turner and M. Benatar, "Ensuring continued progress in biomarkers for amyotrophic lateral sclerosis," *Muscle Nerve*, vol. 51, no. 1, pp. 14–8, 2015. DOI: 10.1002/mus.24470.

[108] N. S. Verber, S. R. Shepheard, M. Sassani, H. E. McDonough, S. A. Moore, J. J. P. Alix, I. D. Wilkinson, T. M. Jenkins, and P. J. Shaw, "Biomarkers in motor neuron disease: A state of the art review," *Frontiers in Neurology*, vol. 10, no. 291, 2019. DOI: 10.3389/fneur.2019.00291. [Online]. Available: https://www.frontiersin.org/article/10.3389/fneur.2019.00291.

[109] T. J. Littlejohns, J. Holliday, L. M. Gibson, S. Garratt, N. Oesingmann, F. Alfaro-Almagro, J. D. Bell, C. Boultwood, R. Collins, M. C. Conroy, N. Crabtree, N. Doherty, A. F. Frangi, N. C. Harvey, P. Leeson, K. L. Miller, S. Neubauer, S. E. Petersen, J. Sellors, S. Sheard, S. M. Smith, C. L. M. Sudlow, P. M. Matthews, and N. E. Allen, "The uk biobank imaging enhancement of 100,000 participants: Rationale, data collection, management and future directions," *Nature Communications*, vol. 11, no. 1, p. 2624, 2020. DOI: 10.1038/s41467-020-15948-9.

[110] N. Weiskopf, J. Suckling, G. Williams, M. M. Correia, B. Inkster, R. Tait, C. Ooi, E. T. Bullmore, and A. Lutti, "Quantitative multi-parameter mapping of r1, pd(*), mt, and r2(*) at 3t: A multi-center validation," *Front Neurosci*, vol. 7, p. 95, 2013. DOI: 10.3389/fnins.2013.00095.

[111] K. Tabelow, E. Balteau, J. Ashburner, M. F. Callaghan, B. Draganski, G. Helms, F. Kherif, T. Leutritz, A. Lutti, C. Phillips, E. Reimer, L. Ruthotto, M. Seif, N. Weiskopf, G. Ziegler, and S. Mohammadi, "Hmri – a toolbox for quantitative mri in neuroscience and clinical research," *NeuroImage*, vol. 194, pp. 191–210, 2019. DOI: 10.1016/j.neuroimage.2019.01.029.

[112] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. P. Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C. N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D. A. Nicholson, D. R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G. A. Price, G.-L. Ingold, G. E. Allen, G. R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J. T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J. V. de Miranda Cardoso, J. Reimer, J. Harrington, J. L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N. J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P. A. Brodtkorb, P. Lee, R. T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, *et al.*, "Scipy 1.0: Fundamental algorithms for scientific computing in python," *Nature Methods*, vol. 17, no. 3, pp. 261–272, 2020. DOI: 10.1038/s41592-019-0686-2.

[113] S. Johnson, *The nlopt nonlinear-optimization package*, Computer Program. [Online]. Available: http://github.com/stevengj/nlopt.