

Efficient Multi-Modal Image Registration based on Gradient Orientations of Minimal Uncertainty

Dante A. De Nigris Moreno



Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

August 2014

A thesis submitted to McGill University in partial fulfillment of the requirements of the
degree of Doctor of Philosophy.

© 2014 Dante A. De Nigris Moreno

THIS PAGE INTENTIONALLY BLANK

Acknowledgments

I would first like to thank my supervisor, Tal Arbel, who provided generous support and supervision throughout my studies. She never failed to thoughtfully review any of the versions of the material found in this thesis, as well as to provide much appreciated funding for all related academic endeavours. I would also like to thank my advisory committee: D. Louis Collins and Jeremy Cooperstock, who provided wise guidance throughout my research efforts. D. Louis Collins offered key insights and feedback throughout my studies. He was also generous to include me in activities and meetings related to his IGNS research group, which proved to be a valuable resource and collaboration for this work.

I would like to thank other professors from CIM whose experience and knowledge were deeply inspiring, such as Kaleem Siddiqi, James Clark, Frank Ferrie, among many others.

I would also like to thank my friends and colleagues at CIM and the MNI, especially Colm Elliott, Meltem Demirkus, Nazanin Aghaloo, Nagesh Subbanna, Mohak Shah, Andrew Jesson, Simon Drouin, Anka Kochanowska and Ian Gerard, for sharing a big part of this path with me. Thanks to Rupert Brooks, his thesis work and guidance served as a valuable reference. I am particularly grateful for Colm's friendship since the very beginning of the Ph.D, for inviting me to his softball team, and for letting me teach Spanish to his daughters, Orla and Olivia.

All graduate students need a strong support network to help them through the ups and downs of thesis writing, and I am certainly incredibly thankful for the uninterrupted encouragement from my friends and family. Thanks to my parents, Myrna and Juan Manuel, for their patience, understanding, guidance, and motivation.

Finally, the most special thanks goes to Emma for sharing so much with me throughout the years, particularly for having the strength and courage of sharing a home with a fellow Ph.D. student. Lo mejor está por venir.

Abstract

This thesis presents a general framework for the registration of medical images across multiple clinical contexts involving rigid and non-rigid applications. The proposed framework relies on gradient orientations as primitive geometric descriptors so as to locally assess image similarity based on orientation alignment and evaluates the metric on sparse locations corresponding to anatomical boundaries of interest. The two main advantages brought forward by the proposed approach are: (1) a substantial reduction in computational complexity and processing time and (2) a significant improvement in robustness against multi-modal contexts with widely different image formation models and significant non-homogeneities.

The proposed approach is evaluated in multiple clinical contexts and compared against state-of-the-art techniques. In the context of neurosurgery, image registration can be employed so as to update a pre-operative magnetic resonance image (MRI) based on an intra-operative ultrasound volume. The proposed approach is evaluated in this challenging time-sensitive scenario and is shown to provide robust performance with sub-second processing times. In the context of the rigid registration of computational tomography (CT) and MRI brain volumes, the proposed approach is evaluated with a publicly available dataset and compared against previously proposed techniques. The quantitative results demonstrate that the proposed approach can employ highly reduced sampling rates (e.g. only 0.05% of the voxels in the image) while still yielding a median registration error inferior to 1mm. In the context of the non-rigid registration of inter-patient MRI brain volumes, the proposed approach is evaluated with a publicly available dataset which measures registration accuracy in terms of the agreement of spatially mapped labels with expert annotated labels. The use of such dataset allows for a fair and unbiased comparison with over fourteen competing techniques. The quantitative results show that the proposed approach achieves slightly inferior accuracy than the top performing method but with only one sixth of the processing time required by the alternative technique. Finally, the proposed approach is evaluated in the context of automatic brain lesion detection which relies on healthy tissue probability maps obtained via registration to a brain atlas. The quantitative comparison against two leading image registration techniques shows that the proposed approach can lead to a slightly improved performance of brain lesion detection algorithms while requiring only one sixth of the processing time used by competing registration approaches.

Sommaire

Cette thèse porte sur de nouvelles techniques de recalage d'images médicales dans différents contextes cliniques et avec applications rigides et non-rigides. Le principe de base est l'utilisation de gradients d'orientation en tant que descripteurs géométriques primaires. Cette technique permet d'évaluer localement la similitude de l'image en se basant sur l'alignement de l'orientation et de restreindre l'évaluation de la mesure de similarité à un ensemble de voxels clairsemés qui correspondent aux points anatomiques d'intérêt. Les deux principaux avantages avancés par l'approche proposée sont les suivants: (1) une réduction substantielle de la complexité de calcul et (2) une amélioration de la robustesse lors du recalage entre différentes modalités d'imagerie.

L'approche est évaluée dans divers contextes cliniques et comparée aux techniques de pointe. Dans le contexte de la neurochirurgie, le recalage d'images peut être employé pour mettre à jour une image obtenue par résonance magnétique (IRM) pré-opératoire en fonction d'une échographie intra-opératoire. L'approche proposée est évaluée dans ce scénario et montre une performance de recalage robuste avec des temps de traitement inférieur à une seconde. Dans le cadre du recalage rigide de la tomographie axiale calculée par ordinateur (TACO) et des volumes du cerveau par IRM, l'approche est évaluée avec un ensemble de données et comparée à des techniques proposées précédemment. Les résultats quantitatifs montrent que l'approche peut utiliser des taux d'échantillonnage très réduits et donne une erreur de recalage médiane inférieure à 1 mm en utilisant seulement 0,05% des voxels. Dans le contexte du recalage non-rigide de volumes du cerveau par IRM, l'approche est évaluée avec un ensemble de données qui mesure la précision de recalage en terme du chevauchement des régions anatomiques. L'utilisation de ces données permet une comparaison impartiale avec plus de quatorze techniques concurrentes. Les résultats montrent que l'approche atteint une précision légèrement inférieure à la méthode la plus performante, mais avec un sixième du temps de traitement requis par la même technique de recalage. Enfin, l'approche proposée est évaluée dans le contexte de la détection automatique des lésions cérébrales qui s'appuie sur des probabilités de tissus sains obtenus par recalage entre l'IRM et un atlas de référence. La comparaison avec deux principales techniques de recalage montre que l'approche proposée peut conduire à une légère amélioration de performance des algorithmes de détection de lésions cérébrales tout en exigeant seulement un sixième du temps de traitement requis par les approches de recalage concurrentes.

Contents

Acknowledgments	i
Abstract	ii
Sommaire	iii
Contents	iv
List of Tables	viii
List of Figures	x
Acronyms	xiii
1 Introduction	1
1.1 Contributions	4
1.1.1 Published Work	7
1.2 Structure of the Thesis	8
2 Background and Previous Work	11
2.1 Overview of Image Registration	12
2.1.1 Ill-posedness and Regularization	13
2.1.2 Localized Image Similarity	14
2.1.3 Feature Extraction	15
2.1.4 Characterization of a Registration Algorithm	16
2.1.5 Transformation Function	17
2.1.6 Similarity Metric	21

2.1.7	Optimization	28
2.1.8	Improvements in Computational Efficiency	30
2.2	Medical Applications	32
2.2.1	Multi-Modal Fusion of Anatomical Images	33
2.2.2	Anatomical Segmentation	34
2.2.3	Longitudinal Analysis	35
2.2.4	Image Guided Interventions	35
2.3	Validation of Registration Techniques	37
2.3.1	Simulated Registration Tasks	37
2.3.2	Distance between Homologous landmarks	37
2.3.3	Agreement between Spatially Mapped Labels	38
2.4	Summary	39
3	Image Registration based on Gradient Orientations of Minimal Uncertainty	42
3.1	Multi-Modal Rigid Registration based on Gradient Orientations	45
3.1.1	Methodology	46
3.1.2	Transformation of Image Gradients	52
3.1.3	Algorithmic Description	56
3.2	Diffeomorphic Registration Based on Gradient Orientations	58
3.2.1	Methodology	59
3.2.2	Diffeomorphism as the Integration of a Velocity Field	62
3.2.3	Transformation of Image Gradients	65
3.2.4	Optimization	70
3.3	Summary	74
4	Image Guided Neurosurgery: Clinical Context and Results	77
4.1	Intra-cranial Tumour Resection	78
4.2	Neuronavigation	79
4.3	Navigation with Tracked Ultrasound	83
4.3.1	Navigation Accuracy	86
4.3.2	Registration of Brain MRI to US	88
4.4	Experiments	92

4.4.1	Clinical Dataset	92
4.4.2	Validation Strategy	92
4.4.3	Results	93
4.5	Inter-Expert Variability	101
4.6	Impact	107
4.6.1	IBIS NeuroNav	107
4.7	Conclusions	110
5	Rigid Registration of CT to MRI Brain Volumes	113
5.1	Clinical Applications	114
5.2	Retrospective Image Registration Evaluation Project	114
5.3	Methodology	115
5.4	Results	116
5.4.1	Random Sub-Voxel Location Sampling	116
5.4.2	Sampling Rate	117
5.4.3	Sensitivity to Magnitude Threshold	118
5.4.4	Performance Comparison with Alternative Approaches	121
5.5	Conclusions	126
6	Automatic Segmentation based on the Registration of Inter-Patient MRI Brain Volumes	128
6.1	Clinical Data	130
6.2	Experimental Results	131
6.2.1	Sensitivity to Spatial and Temporal Discretization	131
6.2.2	Comparison to other methods	132
6.2.3	Comparison with Alternative Edge Detection	137
6.2.4	Sensitivity to Sampling Rate	139
6.3	Conclusions	142
7	Tissue Priors via the Registration of Patient to Template MRI Brain Volumes in the Presence of Pathology	145
7.1	Multiple Sclerosis Brain Lesion Detection And Segmentation	146
7.2	Clinical Data	151
7.2.1	T2 Lesion Ground Truth	153

7.2.2	Gad Lesion Ground Truth	153
7.3	Results	153
7.3.1	Gad Lesions	154
7.3.2	T2 Lesions	156
7.4	Conclusions	157
8	Conclusions and Future Directions	160
8.1	Discussion and Future Directions	163
8.1.1	Edge Detection	164
8.1.2	Regularization and Parameterization of the Transformation Space .	165
8.1.3	Context Learning	166
8.1.4	Validation	167
8.1.5	Other Extensions	168
A	Mathematical Glossary	171
References		173

List of Tables

3.1	Comparison of processing time for evaluating transform image gradients.	56
4.1	Rigid registration results with proposed method evaluated as the mean distance between all the landmarks independently identified by all experts combined.	96
4.2	Statistical summary of rigid registration results with all evaluated techniques.	100
4.3	Analysis of homologous landmarks identified by Expert 1.	102
4.4	Analysis of homologous landmarks identified by Expert 2.	104
4.5	Analysis of homologous landmarks identified by Expert 3.	105
5.1	Rigid registration results expressed in median registration error (in mm) of the VOIs of all cases found in RIRE dataset.	121
5.2	Rigid Registration Results expressed in median registration error (in mm) of the VOIs of all cases found in RIRE dataset.	122
5.3	Rigid registration results expressed in median error (in mm) of the VOIs of all cases found in the RIRE Dataset.	125
5.4	Identification Numbers for RIRE website corresponding to evaluations of the proposed approach configured with $N = 32$	126
6.1	Registration performance and processing time for each registration stage with a static velocity field	135
6.2	Registration performance and processing time for each registration stage with a dynamic velocity field	135

- 6.3 Registration performance and processing time for each registration stage
with a static velocity field and two alternative edge detection techniques. . . 139

List of Figures

2.1	Uniform grid of B-spline control points.	19
2.2	Diffeomorphism and corresponding inverse.	21
2.3	Diffeomorphism characterized as the integration of a velocity field.	22
2.4	MRI vervet brain volume affected by a bias field.	25
2.5	Multi-resolution image pyramid.	31
2.6	Case involving the rigid registration of brain CT and MRI volumes.	33
2.7	Publicly available datasets of expert labelled brain MRI volumes used for the evaluation of non-linear registration techniques.	40
3.1	Algorithmic diagram of proposed rigid registration method.	47
3.2	Posterior density of image gradient orientation.	51
3.3	Multi-scale pixel selection based on gradient magnitude threshold.	53
3.4	Inversely Transformed Gradient Orientation.	54
3.5	Detected edges on a T1-weighted MRI brain volume.	60
3.6	Algorithmic outline of diffeomorphic registration technique.	61
4.1	Neuronavigation with tracked 3D pointer.	80
4.2	Flowchart of a surgical procedure based on neuronavigation.	81
4.3	Landmark based patient to volume registration	81
4.4	Brain-shift caused by changes in intra-cranial pressure.	82
4.5	Illustration of brain-shift with an intra-operative MRI.	83
4.6	Neuronavigation with tracked ultrasound.	84
4.7	Tracked ultrasound probe	85
4.8	Commercial neuronavigation system with support for US.	86
4.9	Operating room configuration integrating an intra-operative MRI.	87

4.10	Pre-operative MRI and iUS in a brain tumour resection surgery.	90
4.11	Pre-operative MRI and iUS slices from Case 3 before and after registration.	95
4.12	Rigid registration results with proposed method evaluated the set of all landmarks independently identified by three different experts.	97
4.13	Rigid registration results with proposed method evaluated with each of three sets of landmarks identified by a particular expert.	98
4.14	Homologous landmarks identified by Expert 1 and 2 for Case 1.	103
4.15	Homologous landmarks identified by Expert 1 and 2 for Case 13.	106
4.16	GPU data allocation for evaluating the similarity function in parallel. . . .	109
5.1	Median mVRE using all voxels within a sampling mask defined by the top 10% voxels with highest gradient magnitude.	117
5.2	Median mVRE evaluated with varying values of N and using randomly selected subsets of voxels.	119
5.3	Sampling masks obtained from top percentage of high gradient magnitude voxels.	120
5.4	Median mVRE evaluated with different sampling masks obtained from varying thresholds and using randomly selected subsets of voxel within each mask.	120
5.5	Box plots showing the VOI registration error of 20 different cases of CT to rectified MRI rigid registration with varying number of voxels.	124
6.1	Expert segmented brain regions from the first case of the CUMC12 dataset [84].	130
6.2	Comparison of segmentation overlap and processing time for a static velocity field with varying values of velocity voxel spacing and number of time steps.	133
6.3	Comparison of segmentation overlap and processing time for a dynamic velocity field with varying values of velocity voxel spacing and number of time steps.	134
6.4	Compound mean target and union (Jaccard) overlap across regions and across cases.	136
6.5	Mean target overlap and standard deviation per region.	138
6.6	Registration performance and processing time with different sampling rates evaluated with a static velocity field.	140
6.7	Registration performance and processing time with different sampling rates evaluated with a dynamic velocity field.	141

7.1	ICBM 2009a Brain Atlas.	148
7.2	Expert labelled brain lesion overlayed onto alternative WM priors.	149
7.3	Gad-enhanced lesions.	150
7.4	T2 lesions.	152
7.5	ROC-like curve illustrating the trade-off between sensitivity and FDR for a Gad lesion detection method with varying healthy tissue priors.	154
7.6	Gad lesion detection performance across lesion size. ¹	155
7.7	T2 lesion detection performance for different lesion sizes.	156
7.8	Example of a medical image analysis processing pipeline.	157

Acronyms

ANTs	Advanced Normalizations Tools
CUMC12	Columbia University Medical Center Dataset (12 cases)
CT	Computer Tomography
dMRI	Diffusion Magnetic Resonance Imaging
DTI	Diffusion Tensor Imaging
EM	Expectation Maximization
GOA	Gradient Orientation Alignment
ICBM	International Consortium for Brain Mapping
IGNS	Image Guided Neurosurgery
IGRT	Image Guided Radiation Therapy
IGRS	Image Guided Radiosurgery
LMI	Local Mutual Information
MALF	Multi-Atlas Label Fusion
MI	Mutual Information
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
NCC	Normalized Cross Correlation
NLSD	Non-Local Shape Descriptor
NMI	Normalized Mutual Information
NRCC	Normalized Regional Cross Correlation
OCT	Optical Coherence Tomography
PET	Positron Emission Tomography
RRMS	Relapsing Remitting Multiple Sclerosis

RIRE	Retrospective Image Registration Evaluation
RGB	Red-Green-Blue Color Model
SyN	Symmetric Image Normalization Method
TPS	Thin-Plate Splines
US	Ultrasound; Ultrasonography

Chapter 1

Introduction

One of the most critical and recurring tasks encountered in a variety of medical image analysis contexts involves finding the spatial mapping between two images such that corresponding pixels share a common underlying anatomical location. Such inference problem is commonly known as image registration and involves challenges which vary significantly with each particular clinical context. For example, some contexts involve the registration of images brought forward by different imaging technologies (i.e. modalities). Such multi-modal contexts face the critical challenge of robustly and accurately identifying (and quantifying) if overlapping image regions correspond to the same underlying location even though the spatially corresponding pixel pairs typically have non-matching intensity values, the modalities involved may not expose exactly the same anatomical information, and each modality involves a unique set of imaging artefacts. Some contexts attempt to infer a complex physical deformation characterized with an extremely large number of degrees of freedom (potentially in the order of millions) from two medical images acquired at different time points. Such registration scenarios face the critical challenge of efficiently solving an ill-posed problem, with a non-unique solution, which involves an extremely large number of unknowns.

The wide-ranging diversity of medical imaging contexts which rely on image registration has led to the proposal of a similarly broad number of image registration techniques. Nonetheless, a significant number of the most popular and successful state-of-the-art registration techniques rely on the common principle of defining an image similarity metric which assesses the similarity of two images throughout their entire domain. Conventional

methods have relied on similarity metrics, such as cross correlation and mutual information, which are defined as a function of the corresponding pixel intensity values. However, pixel intensity values are directly coupled with pixel intensity artefacts specific to each image modality, such as noise, occlusion, and a non-homogeneous intensity response, which can quickly break down the underlying assumption of intensity-based similarity metrics regarding a fixed relationship between corresponding image intensities. Some recent methods have thus proposed extending the image feature space from single pixel values to a set of values which characterizes the pixel-centric image neighbourhood, so as to define a more robust measure of image similarity based on a more descriptive set of image features per location. Unfortunately, such strategies almost inevitably lead to a significant increase in computational complexity. In particular, techniques which rely on a globally evaluated metric must perform a series of computational operations on each of the millions of pixels commonly found in a 3D medical volume. Moreover, in techniques involving an extended image feature space, the computational operations involved at each pixel involve a substantial computational cost, such as the one related to evaluating the set of transformed image features found after deforming the image with a spatial transformation.

There are many clinical contexts that simply cannot incur an increase in computational complexity and processing time, but are nonetheless faced by widely different image modalities which do compel the use of registration frameworks that go beyond conventional intensity-based similarity metrics. For example, in the context of image-guided neurosurgery (IGNS), a pre-operative magnetic resonance image (MRI) may be registered to an intra-operative ultrasound (US) so as to provide an updated MRI volume which accurately reflects the current anatomy and allows the clinician to minimize damage on eloquent tissue. This constitutes a very challenging registration problem where the difficulty largely lies in the radically distinct nature of the modalities involved and the need for significantly reduced processing times that avoid any additional time overhead. Specifically, US images indicate changes in acoustical impedance in the direction of the sound wave, and therefore expose tissue boundaries and liquid-filled structures relatively well. However, they are affected by complicated imaging artefacts like speckle, non-Gaussian noise, shadows and depth sensitivity. On the other hand, MRI volumes can expose different tissues very well and are characterized by a relatively simple noise model. Common image registration methods have great difficulty addressing this context since they typically rely on a pixel-intensity correspondence assumption, which is easily broken by the complex

non-homogeneous imaging artefacts found in US. Furthermore, such methods typically rely on the dense evaluation of a similarity metric across the full spatial domain, which involves non-negligible processing times.

The work presented in this thesis aims to directly address these limitations and proposes a general registration framework which achieves a substantial reduction in processing times, as well as a significant improvement in robustness in challenging multi-modal scenarios. In particular, this thesis argues that it is possible and more reasonable to seek a *structural* correspondence between modalities, where image gradient orientations are used as primitive geometric descriptors for the underlying scene's structure and an optimal spatial mapping is one that optimally aligns spatially corresponding gradients. There are two critically important advantages of using gradient orientations instead of voxel intensities. The first advantage is that gradient orientation alignment is a purely geometric measure of image similarity which is strongly decoupled from image artefacts affecting voxel intensities. This provides an important advantage in terms of robustness against complex and non-homogeneous image formation models, where gradient orientations of prudently selected locations continue to provide an accurate local representation of structures found in the underlying scene. The second advantage is that the locations used for evaluating gradient orientation alignment should intuitively correspond to locations of true underlying boundaries which are shown to correspond to high gradient magnitude locations and far less numerous than the number of voxel locations found in the entire image domain. In other words, the use of gradient orientations leads to a local similarity metric evaluated on a highly reduced set of locations, as opposed to a metric that is evaluated over the entire image domain, which consequently yields a significant reduction in processing times.

The advantages brought forward by the proposed techniques are critically important for the successful integration of intra-operative US in IGNS. This thesis includes a series of quantitative experimental results on real neurosurgical clinical cases, which illustrate how a robust registration performance can be obtained in short processing times. The generality of the proposed approach is further demonstrated with quantitative results that illustrate important performance gains in a variety of medical image registration scenarios. The following section outlines the specific contributions that are present in this work.

1.1 Contributions

The main philosophy of this thesis is that by relying on image gradient orientations (as opposed to voxel intensities) on a sparse set of prudently selected locations, one can propose a robust image registration technique which generalizes to various multi-modal settings and yields highly reduced processing times. In particular, the proposed registration framework is shown to provide important performance gains in the rigid registration of pre-operative MRI to intra-operative US for neurosurgical procedures, the rigid registration of CT and MRI brain volumes, the inter-patient non-rigid registration of MRI brain volumes for the automatic labelling of anatomical regions, and the patient to atlas non-rigid registration of MRI brains volumes for the generation of healthy tissue probability maps. The contributions relating to the development of the registration framework and its contributions to different clinical contexts are enlisted below:

- **A rigid registration method which incorporates a selective gradient orientation alignment metric for improved registration accuracy with drastically reduced sampling rates**

This thesis presents a highly efficient multi-modal rigid registration technique which involves the optimization of a gradient orientation alignment metric over a sparse set of locations involving a reduced gradient orientation uncertainty. The proposed method is the first to demonstrate that superior accuracy can be achieved by incorporating a sub-voxel location sampling strategy (i.e. sampled locations are not restricted to voxel grid locations)¹ and by increasing the selectivity of the gradient orientation alignment metric so that only tightly aligned gradient orientations are rewarded. Furthermore, it also allows the use of highly aggressive sampling rates (i.e. number of evaluated voxels) while continuing to yield sub-millimeter accuracy.

- **The first system to permit sufficiently fast and accurate registration of intra-operative ultrasound (US) to pre-operative MRI to be integrated into the operating room in the context of image-guided neurosurgery**

The work in this thesis yields concrete results towards improving the integration of

¹Previous work related to sub-voxel location sampling for image registration has been generally limited to illustrate how it affects registration robustness in specific contexts and is also generally limited to voxel-intensity based similarity metrics.

iUS in the operating room for the purpose of improved guidance throughout neurosurgery. In particular, the implementation of the proposed technique in a GPU provides the first rigid registration framework which can robustly address the registration of MRI to iUS with processing times inferior to one second, where competing techniques involve processing times in the order of minutes. Such performance, in conjunction with a fast GPU-based volume reconstruction technique (also implemented as part of this thesis), provides a minimal turnaround time for the surgeon and substantially reduces the overall delay to obtain updated pre-operative data. Such contributions were implemented so as to extend a custom neuronavigation framework and are now part of standard protocol for using iUS in neurosurgery in ongoing research projects at the Montreal Neurological Institute.

- **A novel diffeomorphic registration framework based on the symmetric consistency of inferred boundaries evaluated with gradient orientation alignment**

The proposed diffeomorphic registration approach seeks the optimal alignment of inferred boundaries in a symmetric fashion. Additionally, it involves an initial inference of boundaries, which defines a sparse set of locations over which a local similarity metric is evaluated. A diffeomorphic transformation is then optimized over the set of boundary locations at both images. As such, it is the first non-rigid registration approach that combines the notions of gradient orientation alignment and point-to-point distance in a symmetrically consistent framework. The proposed approach is evaluated in the context of the registration of inter-patient MRI brain volumes and the generation of healthy brain tissue priors via the registration to a standard template. The context of inter-patient MRI brain volumes is evaluated with a commonly used publicly available dataset and results demonstrate that the proposed technique can achieve an accuracy similar to top ranking state-of-the-art methods, but with significantly reduced processing times. In particular, the proposed registration framework involves average processing times ranging between 1 and 6 minutes, as opposed to processing times of around half an hour with alternative approaches. The context involving the generation of healthy brain tissue priors via image registration was evaluated in terms of how different tissue priors affect the performance of subsequent lesion detection algorithms. The corresponding results illustrate how the tissue priors

generated with the proposed approach can be obtained in significantly reduced processing times and can lead to improved detection performance in two different brain lesion detection algorithms.

- **The first multi-scale voxel selection strategy based on gradient orientations of minimal uncertainty**

The theoretical framework presents the first analytical multi-scale framework for characterizing the effect of additive Gaussian noise on computed image gradient orientations. The task is defined as a Bayesian inference problem, where the computed image gradient magnitude and orientation constitute “observations” which shape the posterior distribution of the “true” underlying gradient orientation. This theoretical exercise provides important conclusions that allow for principled design choices in the proposed registration algorithm. In particular, the analysis demonstrates that the uncertainty of the gradient orientation is an inversely monotonic function of the observed gradient magnitude. In other words, locations with high gradient magnitude constitute locations with reduced gradient orientation uncertainty. It also demonstrates that the uncertainty of gradient orientations can be directly compared across scales. This allows the user to define a unique criterion for selecting gradient orientation locations at all scales at which the gradient is computed.

- **The first rigid registration framework which incorporates the pullback differential for efficient evaluation of transformed gradients**

The work in this thesis provides a general framework for reducing the computational complexity of transformation gradient orientations in the context of image registration. In particular, it emphasizes that the transformation of image features based on first-order differentials analytically involves the use of the pullback differential. In simple terms, the evaluation of an image gradient at a given location after a transformation additionally involves the pre-multiplication with the Jacobian of the transformation at such point. This property is particularly powerful in the context of rigid registrations with a sparse set of evaluated locations, where the Jacobian matrix is constant at all locations and the registration algorithm will hence involve a minimal computational overhead.

Ultimately, the variety of positive results presented in this thesis provide a clear indication that the use of image gradient orientations on a sparse set of prudently selected locations for image registration can successfully reduce the processing time of rigid and non-rigid algorithms by significant factors, while maintaining a highly accurate registration performance across multiple contexts.

1.1.1 Published Work

The publications related to the work presented in this thesis can be found below.

Peer-Reviewed Journal Publications

- [45] D. De Nigris, D. L. Collins, and T. Arbel. Multi-Modal Image Registration based on Gradient Orientations of Minimal Uncertainty. *IEEE Transactions on Medical Imaging*
- [47] D. De Nigris, D. L. Collins, and T. Arbel. Fast Rigid Registration of Pre-Operative Magnetic Resonance Images to Intra-Operative Ultrasound for Neurosurgery based on High Confidence Gradient Orientations. *International Journal of Computer Assisted Radiology and Surgery*

Peer-Reviewed Conference Publications

- [44] D. De Nigris, D. L. Collins, and T. Arbel. Fast and Robust Registration Based on Gradient Orientations: Case Study Matching Intra-Operative Ultrasound to Pre-Operative MRI in Neurosurgery. *Information Processing in Computer-Assisted Interventions*
- [43] D. De Nigris, D. L. Collins, and T. Arbel. SymBA: Diffeomorphic Registration Based on Gradient Orientation Alignment and Boundary Proximity of Sparsely Selected Voxels In *6th International Workshop on Biomedical Image Registration*

Related Work

The work in the following publications does not appear in this thesis, but is related to this research.

- [48] D. De Nigris, L. Mercier, R. Del Maestro, D. L. Collins, and T. Arbel. Hierarchical Multimodal Image Registration Based on Adaptive Local Mutual Information. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2010*
- [46] D. De Nigris, D. L. Collins, and T. Arbel. Deformable Registration of Chest CT Scans with Adaptive Local Mutual Information *Grand Challenges in Medical Image Analysis - MICCAI 2010*

- [110] K. Murphy, et al. Evaluation of Methods for Pulmonary Image Registration: The EMPIRE10 Study In *Grand Challenges in Medical Image Analysis - MICCAI 2010*

1.2 Structure of the Thesis

The remainder of this document is organized as follows.

The field of image registration is quite vast and varied, in terms of both techniques and applications. Chapter 2 provides a brief high-level overview of the taxonomy of image registration algorithms, a description of a few key medical applications, and a discussion regarding the challenge of choosing a suitable validation strategy for image registration techniques. Even though such overview is not exhaustive, it provides an important reference for identifying how the work of this thesis relates to other image registration algorithms.

Chapter 3 presents the theoretical and technical details involving the registration techniques proposed. Section 3.1.1 presents the analysis for characterizing the uncertainty of gradient orientations in a multi-scale framework. Section 3.1.2 demonstrates the use of the pullback differential for evaluating the gradient orientation of a transformed image. Based on those results, Section 3.1.3 defines a rigid registration technique, which is materialized in both a CPU and a GPU-based implementation. Section 3.2 presents the proposed non-rigid registration technique involving a similarity metric based on gradient orientation alignment and distance to the nearest boundary. A significant effort is devoted to explain the use of a time-constant or time-dependent velocity field to characterize a diffeomorphism. In particular, it derives the similarity metric gradient expressions that result from adopting a velocity field with a coarse spatial discretization which explicitly accounts for the dependency to an interpolation function.

Chapter 4 highlights the relevance of the proposed registration techniques in the context of IGNS. Section 4.1 provides a brief overview of the clinical background regarding brain tumour resection and its need for image guidance. Section 4.2 describes the technical setup involved in a common neuronavigation context and the need for incorporating intra-operative imaging. Section 4.3 explains the appeal of adopting intra-operative US (iUS) in such procedures, and the technical requirements involved in integrating iUS in the operating room, as well as the challenge of registering a pre-operative MRI volume to an iUS image so as to update pre-operative data. Section 4.4 presents a series of results evaluated with a publicly available dataset of real brain tumour resection cases with corresponding pre-operative MRI and iUS. The results obtained with the GPU-based implementation show

that a robust performance can be obtained with sub-second processing times, and is a clear indication of how it can improve the integration of iUS in the procedure. Section 4.5 performs an in-depth analysis of the inter-rater variability based on the expert landmarks found in the publicly available dataset, which illustrates the practical challenges of designing a validation strategy that is consistent, accurate and unbiased. Section 4.6 provides further detail on the technical contributions that were performed as part of this work so as to extend the functionality of an existing neuronavigation framework and which have lead to a more seamless integration of iUS in ongoing clinical procedures. The particular details describing the technical measures that had to be adopted so as to truly profit from the massively parallel computational power of a GPU are described in Section 4.6.1.

Chapter 5 presents an exhaustive set of results in the context of the rigid registration of CT to MRI brain volumes. The experiments relied on a commonly used publicly available dataset, and thus allowed for a direct comparison with previously reported results. The set of quantitative results obtained provide a series of critically important conclusions. Section 5.4.1 shows that in the context of anisotropic volumes, improvements in registration accuracy can be obtained by sampling sub-voxel locations, as opposed to locations restricted to voxel-grid locations. The experiments found in Section 5.4.1 and 5.4.2 demonstrate that the registration accuracy can be further improved by adopting a selective gradient orientation alignment metric where only tightly aligned gradient orientations are considered a match. Section 5.4.3 shows that relying on the gradient magnitude as an indicator on the confidence on such features allows us to select a reduced set of locations from the top percentage of locations with highest gradient magnitude, while maintaining and even improving registration accuracy. The experiments in Section 5.4.3 also demonstrate that the registration technique is not particularly sensitive to the choice of gradient magnitude threshold. Section 5.4.4 then presents a comprehensive comparison of the registration performance of the proposed technique against the performance of multiple previously proposed approaches. The results from such experiments demonstrate that the proposed technique can rely on an extremely low rate of sampled locations to drive the registration technique while still obtaining a sub-voxel accuracy.

Chapter 6 and 7 present results related to the proposed non-rigid registration framework in two different contexts. Chapter 6 involves the context of inter-patient MRI brain volume registration which is evaluated with a commonly used publicly available dataset. In such context, registration performance is evaluated as the overlap of estimated segmentation

labels with expert segmentation labels, and where the estimated labels are obtained from warping the expert labels found in a second volume. Section 6.2.1 presents a comprehensive set of experiments which characterize the registration performance in terms of the multiple parameters involved. Section 6.2.2 compares the registration performance obtained with the proposed approach against the results obtained with 14 alternative state-of-the-art techniques. The comparison shows that the proposed technique yields slightly inferior accuracy than the top ranking methods, but with highly reduced processing times.

Chapter 7 involves the generation of healthy brain tissue priors via a patient to template non-rigid registration. In particular, the clinical context involves the detection and segmentation of multiple sclerosis (MS) brain lesions with automatic algorithms that rely on healthy brain tissue probabilities. Specifically, MS brain lesions involve intensity profiles that have a significant overlap with the profile of healthy tissue types and there is a fundamental need for additional information, such as healthy tissue probability maps, for robustly resolving such lesions. A common approach for obtaining healthy tissue probability maps, also referred to as healthy tissue priors, involves non-linearly registering the patient’s MRI volume to a standard brain atlas that includes healthy tissue probabilities for each voxel. The patient’s healthy tissue priors can be then be directly obtained by evaluating the probability values at the spatially corresponding location in the atlas. The experimental setup found in Chapter 7 evaluates how tissue priors obtained from competing registration techniques affect the performance of two automatic brain lesion detection tasks: the detection of active (i.e. gadolinium-enhanced) MS lesions and the detection of chronic MS lesions. As such, it constitutes a novel approach for validating a registration strategy with an explicit focus on bottom-line performance of the task at hand. The reported results show that both lesion detection algorithms yield a slightly superior detection performance than the two competing state-of-the-art methods. More importantly, the average processing time of the proposed registration strategy is drastically inferior to that of the competing methods, and provides an important advantage for improving the throughput of large scale clinical trials.

Finally, Chapter 8 summarizes a series of conclusions obtained from this work and proposes multiple extensions and variants that could be pursued in the future.

Chapter 2

Background and Previous Work

Image registration concerns the general task of mapping the coordinate space of one dataset to the coordinate space of another dataset. It is an inference problem that appears in various applications in remote sensing, computer vision and medical imaging. For example, stitched panoramic image creation [159] is a popular application which involves the construction of a wide-view panoramic picture from a set of smaller overlapping pictures. It can also be extended to render fully immersive 360° panoramic pictures which can be explored in any direction. Image registration plays a fundamental role in panoramic image creation and various improvements in computational efficiency have been specifically developed so as to allow robust real time performance in modern smartphones.

Image registration also plays an essential role in a variety of medical image analysis tasks involving either a single image modality (mono-modal) or various image modalities (multi-modal). Note that an image modality refers to a specific imaging technology, such as MRI, PET, CT, US, and X-rays. For example, rigid multi-modal registration is commonly used to visualize and analyze various medical image modalities in a common coordinate frame. Non-rigid mono-modal registration techniques are also commonly used for inferring the physical deformation occurring between time points. Such feature may be used for tracking a region of interest (e.g. tumour), characterizing abnormal deformations (e.g. identifying grey matter atrophy), among many other applications.

The image registration field is vast in both applications and proposed techniques. The wide variety of applied contexts with significantly diverse challenges limits the generality of any specific registration technique. Hence, registration frameworks are typically proposed

with a specific application in mind. This chapter provides a brief formal overview of image registration algorithms, including a brief description of the main components characterizing a majority of techniques and how they relate to particular challenges. For a complete and exhaustive survey of image registration methods, the reader is invited to refer to [23, 98, 176].

The chapter is structured as follows. Section 2.1 provides a brief overview of components involved in an image registration framework. Section 2.2 describes the medical contexts where image registration plays an important role and provides an overview of the techniques proposed for such applications. The validation of image registration techniques in real medical contexts is an important challenge which has lead to many proposed strategies with varying advantages and disadvantages. In particular, given the absence of an absolute ground truth, many proposed validation strategies generally involve a limited accuracy in evaluating geometric error, and can only provide a surrogate measure or an estimation of the geometric error. Section 2.3 highlights the challenges in designing a validation strategy that accurately evaluates the performance of an image registration technique in a real medical context and gives a brief overview of commonly used strategies.

2.1 Overview of Image Registration

Image registration concerns the task of finding a spatial transformation function, \mathbf{T} , that brings a moving image, I_m , in spatial correspondence with a fixed image, I_f , so that spatially mapped pixel-pairs relate to a common location in the underlying scene. With an increased effort in providing a general definition, one can formally express such task as an optimization problem of the following form,

$$\mathbf{T}^* = \underset{\mathbf{T} \in \mathcal{T}}{\operatorname{argmax}} E_g(I_f, I_m, \mathbf{T}) \quad (2.1)$$

where $\mathbf{T} : \Omega \rightarrow \Omega$ is a transformation function that maps the coordinate space, Ω , of the fixed image, I_f , to the coordinate space of the moving image, I_m . An image, I , corresponds to a function that maps the coordinate space to n scalar values, $I : \Omega \rightarrow \mathbb{R}^n$. For example, a grey-scale image has one scalar value per location (i.e. $n = 1$) and a RGB image has three scalar values per location (i.e. $n = 3$). Finally, E_g is an energy functional dependent on I_f , I_m and \mathbf{T} , which yields a maximal value when the images are aligned.

2.1.1 Ill-posedness and Regularization

Many of the contexts encountered by image registration techniques are characterized by image data that does not lead to a single solution but to multiple equally valid solutions. In other words, the image data may not be sufficient to restrict the inference problem to a single solution. Such problems are ill-posed and a common strategy to address them has been to rely on regularization.

The regularization strategies encountered in image registration techniques [32, 53, 64, 133, 137, 155] are typically characterized by incorporating a penalty term in the objective function,

$$E_g(I_f, I_m, \mathbf{T}) = S_g(I_f, I_m \circ \mathbf{T}) - \mathcal{R}(\mathbf{T}) \quad (2.2)$$

where the penalty term, $\mathcal{R} : \mathcal{T} \rightarrow \mathbb{R}$, maps a transformation function to a value assessing its improbability and S_g measures the similarity of the fixed image with the transformed moving image. For example, one could assume that a transformation function involving large displacements is less probable than one involving small displacements.

The penalty term is commonly based on the first or second-order derivatives of the transformation function. As such, it is based on the basic assumption that transformations with large local variations are less plausible than transformation with small local variations. For example, in [137] the authors adopt a penalty term based on the second-order derivatives of the transformation function to regularize the non-rigid registration of 3D volumes. The penalty term is expressed as

$$\begin{aligned} \mathcal{R}(\mathbf{T}) = & \frac{1}{\Omega} \int_x \int_y \int_z \left[\left(\frac{\partial^2 \mathbf{T}}{\partial x^2} \right)^2 + \left(\frac{\partial^2 \mathbf{T}}{\partial y^2} \right)^2 + \left(\frac{\partial^2 \mathbf{T}}{\partial z^2} \right)^2 \right. \\ & \left. + 2 \left(\frac{\partial^2 \mathbf{T}}{\partial xy} \right)^2 + 2 \left(\frac{\partial^2 \mathbf{T}}{\partial xz} \right)^2 + 2 \left(\frac{\partial^2 \mathbf{T}}{\partial yz} \right)^2 \right] dx dy dz \end{aligned}$$

and corresponds to the 3D counterpart of the 2D bending energy of a thin-plate of metal [168].

Bayesian Interpretation The use of a regularization strategy can be associated to a Bayesian interpretation of the image registration problem, such as the one found in [130], where the objective is to find the transformation that maximizes a posterior probability

$$\mathbf{T}^* = \underset{\mathbf{T} \in \mathcal{T}}{\operatorname{argmax}} p(\mathbf{T} \mid I_f, I_m) \quad (2.3)$$

where the posterior probability of the transformation function may be expressed as,

$$\begin{aligned} p(\mathbf{T} \mid I_f, I_m) &= \frac{p(I_m \mid I_f, \mathbf{T}) \cdot p(\mathbf{T})}{p(I_m)} \\ &\propto p(I_m \mid I_f, \mathbf{T}) \cdot p(\mathbf{T}) \end{aligned}$$

where $p(I_m \mid I_f, \mathbf{T})$ is the likelihood of \mathbf{T} given I_m (conditioned to I_f), and $p(\mathbf{T})$ is the prior density on the transformation function which characterizes the plausibility of \mathbf{T} *independently* of the image data.

In the case where all transformations are equally plausibly, the prior density $p(\mathbf{T})$ is defined by a uniform distribution and the maximum a posteriori (MAP) estimate is equivalent to the maximum likelihood estimate (MLE). In other words, the solution is driven exclusively by the image data. Alternative, if there is a non-uniform prior knowledge regarding the plausibility of each transformation in the set of valid transformations, one can embed such information in the form of a non-uniform prior distribution. Finally, if the optimization is re-expressed in terms of the *log posterior*, one obtains,

$$\mathbf{T}^* = \underset{\mathbf{T} \in \mathcal{T}}{\operatorname{argmax}} \log p(I_m \mid I_f, \mathbf{T}) + \log p(\mathbf{T}) \quad (2.4)$$

where the first term assesses image similarity and the second term assesses the plausibility of the transformation function. Note that the two terms have a direct correspondence with the terms in Equation 2.2.

2.1.2 Localized Image Similarity

Many commonly used similarity functionals are defined as the integration of a *local* similarity functional, S_l , over the coordinate space,¹

$$S_g(I_f, I_m^\downarrow) = \frac{1}{|\Omega|} \int_{\mathbf{x} \in \Omega} S_l(I_f, I_m^\downarrow)(\mathbf{x}) d\mathbf{x}. \quad (2.5)$$

Note that the notation convention, $I_m^\downarrow = I_m \circ \mathbf{T}$, is adopted to compactly express a transformed moving image.

¹Note that in a registration context, the spatial domain, Ω , over which one can evaluate the definite integral will be limited to the overlap of the spatial domain of both images, and one may include a normalization with respect to the size of the spatial domain.

Equation 2.5 may also be further simplified into a local similarity function dependent solely on the image values at the location of interest:

$$S_g(I_f, I_m^\downarrow) = \frac{1}{|\Omega|} \int_{\mathbf{x} \in \Omega} S_l(I_f(\mathbf{x}), I_m^\downarrow(\mathbf{x})) d\mathbf{x}. \quad (2.6)$$

The expression allows for a discretization in the coordinate space that is directly related to the discretization of the images being registered,

$$S_g(I_f, I_m^\downarrow) \approx \frac{1}{|\Omega_Z|} \sum_{\mathbf{x} \in \Omega_Z} S_l(I_f(\mathbf{x}), I_m^\downarrow(\mathbf{x})) \quad (2.7)$$

where Ω_Z denotes a set of *discretized* locations identifying the centre of a fixed image voxel. Note that the computational complexity of evaluating the similarity metric is inversely related to the resolution of the images being registered, and high resolution images have an exponentially increased computational complexity. For example, a 3D image with isotropic voxel spacing of $h \times h \times h$ will have a number of voxels that is proportional to h^{-3} and thus will involve a complexity of $O(h^{-3})$ for evaluating the similarity metric.

Many different similarity metrics have been proposed for addressing image registration problems with varying challenges and they constitute a critical component of any technique. Section 2.1.6, provides an overview of commonly used metrics and discuss their limitations.

2.1.3 Feature Extraction

The types of similarity metrics discussed so far have been limited to functions and functionals of image intensity values. However, many registration algorithms employ an intermediate *feature extraction* function ψ which maps the intensity values found in the original images, (I_f, I_m) , to either an alternate pair of images, (Ψ_f, Ψ_m) , where voxel values now correspond to a feature of interest, or to a sparse set of highly discriminable image features (e.g. corners).

In the context where the feature extraction function leads to an alternate pair of feature images, (Ψ_f, Ψ_m) , the corresponding similarity metric is defined as a function of the feature images,

$$S_g(I_f, I_m^\downarrow) = S_g^\psi(\Psi_f, \Psi_m^\downarrow). \quad (2.8)$$

Such approaches can be referred to as dense-feature based, since they involve image

features computed in the entire spatial domain. For example, some approaches may rely on the optimization of the inner product of normalized image gradients [65] in an attempt to focus the registration task to image boundaries. Section 2.1.6 describes various image registration approaches that rely on dense-features and discusses the limitations they address.

Alternatively, in the context where the feature extraction function leads to a sparse set of highly discriminable image features, the image registration problem can longer be characterized as the maximization of a dense objective function and is instead addressed by sparse point correspondences. Such approaches can be computationally efficient when dealing with transformations that can be accurately resolved with a small number of point correspondences, such as a rigid transformation. The two basic steps of such approaches are to first identify a sparse set of point correspondences based on image feature matches and then analytically solve for the transformation that minimizes the distance between such point correspondences.

One of the most successful examples of this type of approach is the use of SIFT² features [95], which can be found in consumer oriented automatic panoramic image stitching applications [24]. In this approach, a sparse set of highly discriminable SIFT features [95] is identified in each image and the features are described based on the distribution of gradient orientations in their neighbourhood.

The use of highly discriminable sparse local image features is a highly active research topic in the general field of computer vision for tasks such as object detection, object classification and image stitching. Furthermore, there has been much work devoted to the development of different features for such tasks. The reader may refer to [107] for an exhaustive performance comparison of local image features. However, in the context of medical image registration, there is a more notable trend towards using dense-feature based approaches or relying directly on image intensity values, and the use of discriminable sparse local features has been rather limited to a few efforts.

2.1.4 Characterization of a Registration Algorithm

The definition of the registration task found in Equation 2.1 allows us to identify major key components such as: 1) E_g , the energy functional, 2) ψ , the image feature extraction

²SIFT stands for Scale-Invariant Feature Transform

function³, 3) \mathcal{T} , the space of transformation functions, 4) Ω , the coordinate space and 5) n , the number of components at each pixel. Further important specifications of a registration algorithm relate to the optimization strategy used to maximize E_g and all other related numerical methods (e.g. interpolation strategy, differentiation strategy, etc.).

The following section briefly describes some of the previously proposed schemes that have seen success in various registration contexts in terms of such key components.

2.1.5 Transformation Function

The space of valid transformation functions, \mathcal{T} , plays a critical role in characterizing the complexity of the optimization found in image registration. For example, so-called parametric transformations are fully defined by a small set of parameters and may be addressed with robust and efficient off-the-shelf optimization techniques (e.g. simplex method). However, some contexts require solving for a highly complex transformation defined by a very large set of parameters (typically proportional to the number of pixels found in the image) and are commonly addressed with a gradient ascent related technique. The next section briefly describes some of the transformation functions that can be commonly found in image registration techniques.

Linear Transformations

Linear transformations are functions that can be expressed as a matrix multiplication and a translational displacement,

$$\mathbf{T}(\mathbf{x}) = \mathbf{Ax} + \mathbf{b} \quad (2.9)$$

where \mathbf{A} is a $d \times d$ matrix and \mathbf{b} is a translational offset or displacement.

A linear transformation can also be expressed with a single matrix, $\mathbf{L} = \begin{pmatrix} \mathbf{A} & \mathbf{b} \\ 0 & 1 \end{pmatrix}$, in homogeneous coordinates, such that,

$$\begin{pmatrix} \mathbf{T}(\mathbf{x}) \\ 1 \end{pmatrix} = \mathbf{L} \times \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \quad (2.10)$$

Linear transformations are global parametric transformations that are compactly represented by a few parameters. They also involve some practical mathematical properties. For

³When there is no feature extraction ψ can be considered equal to the identity function

example, the composition of linear transformations is obtained as the multiplication of corresponding homogeneous matrices and the inversion of a linear transformation is obtained as the inversion of its homogeneous matrix.

Rigid Transformation

A rigid transformation is a specific type of linear transformation which preserves the distance between every pair of points. It can be characterized by a rotation matrix, \mathbf{R} , and a translational displacement, \mathbf{b} , such that,

$$\mathbf{T}(\mathbf{x}) = \mathbf{Rx} + \mathbf{b} \quad (2.11)$$

The rotation in 2D space can be compactly characterized with one rotation angle. However, the rotation in 3D space can be characterized either with three rotation angles (i.e. Euler angles) or with a quaternion (involving four scalar values).

Thin Plate Splines

Thin Plate Splines (TPS) [19, 52] refers to a non-linear transformation model that satisfies a specified set of point correspondences, $T(\mathbf{x}_i) = T_i$, and minimizes a “certain sort of bending energy” [19]. For a 2D domain, the transformation along a given spatial dimension, i , can be expressed as,

$$T_i = \underset{T}{\operatorname{argmin}} \int \left(\frac{\partial^2 T_i(\mathbf{x})}{\partial x_1^2} \right)^2 + 2 \left(\frac{\partial^2 T_i(\mathbf{x})}{\partial x_1 \partial x_2} \right)^2 + \left(\frac{\partial^2 T_i(\mathbf{x})}{\partial x_2^2} \right)^2 d\mathbf{x}. \quad (2.12)$$

The minimization of such energy with N homologous landmarks leads to a solution of the following form (in 2D),

$$T_i(\mathbf{x}) = a_0 + \sum_{i=1}^d a_i \cdot x_i + \sum_{l=1}^N w_l \cdot K(\mathbf{x} - \mathbf{x}_l) \quad (2.13)$$

$$K(\mathbf{x}) = -\|\mathbf{x}\|^2 \cdot \log(\|\mathbf{x}\|^2) \quad (2.14)$$

where K denotes the *radial basis function* or *kernel function* and where the coefficients \mathbf{a}, \mathbf{w} can be obtained by solving a corresponding linear system.

The TPS transformation is an instance of a more general class of transformation functions called *kernel transformations* or *radial basis transformations*, which are defined as the summation of kernel functions centred at each landmark [42, 154].

B-spline Transformations

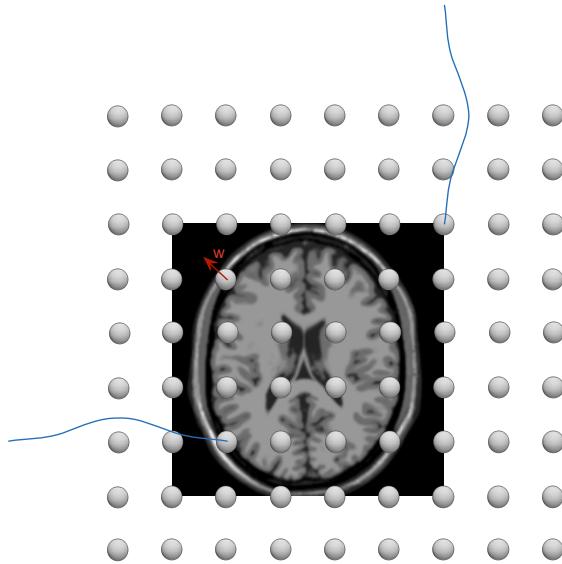


Fig. 2.1 Uniform grid of B-spline control points. Each control point involves a displacement vector, \mathbf{w} , that locally characterizes the deformation field.

B-spline transformations were initially proposed in [137] for the purposes of medical image registration and have since seen significant success in numerous clinical applications [94, 99, 155]. The model defines a non-rigid transformation as the weighted summation of tensor products of cubic B-splines centred on *control points* uniformly distributed in the spatial domain. Fig. 2.1 illustrates a uniform grid of B-spline control points that defines a transformation function within the spatial domain of the image.

For a 2D domain, a B-spline transformation can be expressed as,

$$\mathbf{T}(\mathbf{x}) = \mathbf{x} + \sum_{\mathbf{y} \in \mathcal{N}_{\mathbf{x}}} \mathbf{w}_{\mathbf{y}} \beta_3 \left(\frac{x_1 - y_1}{h_1} \right) \beta_3 \left(\frac{x_2 - y_2}{h_2} \right) \quad (2.15)$$

where β_3 is a third order B-spline function [164], h_i is the spacing between control points along dimension i , \mathbf{y} is the location of a given control point, and $\mathcal{N}_{\mathbf{x}}$ is a neighbourhood of

control points that have a contribution on the transformation of location \mathbf{x} . Note that each tensor product is weighted by a vector \mathbf{w}_y which characterizes the displacement contributed by each control point.

B-spline transformations are very useful for compactly characterizing a non-rigid transformation and provide the flexibility of a user-defined control point grid. In other words, one can choose to adopt a coarse space grid of control points for compactly characterizing a non-rigid deformation with somewhat restricted degrees of freedom. Hence, multi-resolution registration techniques that use B-spline transformations generally adopt a coarse grid of control points at the initial stage and refine the grid of control points at each subsequent registration stage. In other words, each registration stage is composed of images blurred and sampled at a given resolution and a B-spline transformation of a given complexity. An important advantage of gradually increasing the resolution of the control grid is the reduction of computational complexity, since B-spline transformations with coarse grids have a reduced number of transformation parameters.

Diffeomorphisms as the Integration of a Velocity Field

A diffeomorphism [5], φ , is a smooth, bijective transformation function that has a smooth inverse. They were first introduced for computer vision in [163] and have since been developed with multiple variants [10, 16, 33, 108] for medical image registration tasks. There are two major advantages regarding the use of a diffeomorphism in an image registration tasks in comparison with alternative non-linear transformation models such as TPS and B-spline transformations. First, diffeomorphisms explicitly enforce the constraint that there is no tearing or folding of the deformation (given that the transformation and its inverse is smooth and continuous). Second, diffeomorphisms provide both the forward-mapping transformation function and a backward-mapping transformation function. The benefit of having both transformations lies in the ability to map image information in any direction and is illustrated in Fig. 2.2.

In the context of complex registration tasks, diffeomorphisms based on the large deformation diffeomorphic metric mapping (LDDMM) model [16] have enjoyed significant success. In this model, the diffeomorphism is defined as the endpoint, $\varphi = \phi_1$, of the flow

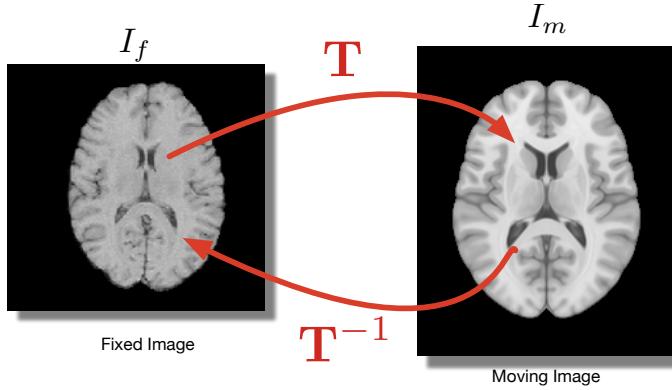


Fig. 2.2 Diffeomorphism and corresponding inverse. A diffeomorphic transformation allows for a smooth and continuous inverse transformation.

of a time-dependent velocity vector field, \mathbf{v}_t

$$\phi_t(\mathbf{x}) = \phi_0(\mathbf{x}) + \int_{u=0}^t \mathbf{v}_u(\phi_u(\mathbf{x})) du \quad (2.16)$$

where $\mathbf{v}_t(\phi_t) = \frac{d\phi_t}{dt}$, $\phi_0 = id$ corresponds to the identity function and \mathbf{x} is a location in the fixed image domain. In other words, φ maps the coordinate space of the fixed image to the coordinate space of the moving image.

The numerical implementation of such approach involves the approximation with a discretization of N time steps such that,

$$\varphi(\mathbf{x}) \approx \mathbf{x} + \sum_{i=1}^N \mathbf{v}_i(\phi_i(\mathbf{x})) \quad (2.17)$$

In other words, the implementation of a diffeomorphism can be alternatively described as the composition of a series of small magnitude displacement fields. Fig. 2.3 illustrates the trajectory of a given location as it advances through the velocity field.

2.1.6 Similarity Metric

The choice of similarity function is critical to the success of image registration and as such needs to take into consideration the nature of the images being registered.

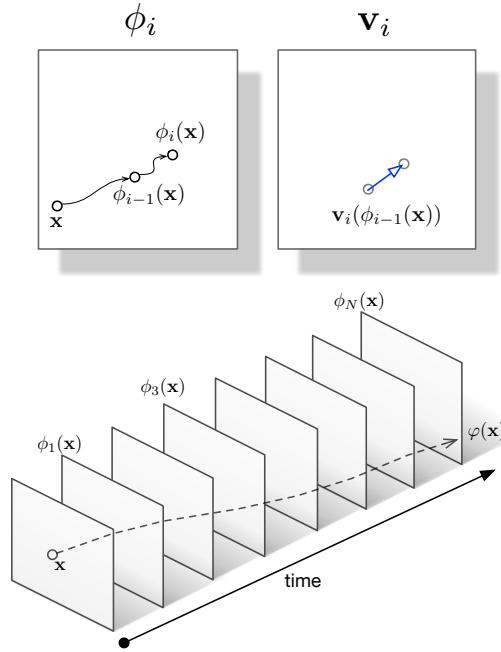


Fig. 2.3 Diffeomorphism characterized as the integration of a velocity field.

Mean Squared Difference

When registering images with a common image formation model, one can assume that spatially corresponding pixels will have similar image intensities. Hence, one can employ a metric which is maximal when the squared difference of the intensities of corresponding pixels is minimal, such as the sum of squared differences (SSD)⁴,

$$S_g^{SSD}(I_f, I_m^\downarrow) = -\frac{1}{|\Omega_Z|} \sum_{\mathbf{x} \in \Omega_Z} (I_f(\mathbf{x}) - I_m^\downarrow(\mathbf{x}))^2. \quad (2.18)$$

The metric was a popular choice in early computer vision techniques [82, 96], and can be found implemented in a variety of medical image registration techniques [58, 66, 162].

One can also obtain the same similarity metric by considering the assumption that the fixed image intensities are related to the moving image intensities by an additive Gaussian noise model [130].

⁴Recall that the objective function is *maximized*, hence the negative sign in the expression for S_g

Cross Correlation

When registering images of different modalities, the task is of increased difficulty. The most obvious challenge is that structures and regions that appear on both modalities will typically have different intensity responses on each modality. For example, voxels corresponding to bone tissue appear with an intensity I_f ('bone') in the fixed image, while appearing with a different intensity I_m ('bone') in the moving image. A possible simplifying assumption is that there is a *fixed* mapping function between inter-modality intensities that can be expressed as a linear function. Under such assumptions, the registration task can be directly associated [130] to the problem of maximizing the cross correlation (CC) or normalized cross correlation (NCC) between corresponding image values,

$$S_g^{NCC}(I_f, I_m^\downarrow) = \sum_{\mathbf{x} \in \Omega_Z} \frac{(I_f(\mathbf{x}) - \bar{I}_f)(I_m^\downarrow(\mathbf{x}) - \bar{I}_m^\downarrow)}{\sigma_{I_f} \sigma_{I_m^\downarrow}} \quad (2.19)$$

where $\bar{I} = \frac{1}{|\Omega_Z|} \sum_{\mathbf{x}} I(\mathbf{x})$ and $\sigma_I^2 = \frac{1}{|\Omega_Z|} \sum_{\mathbf{x}} (I(\mathbf{x}) - \bar{I})^2$.

The metric was initially proposed for medical image registration contexts in [13, 60] and has since been implemented in other works [10, 29].

Mutual Information

There are various contexts whereby a relationship between inter-modal intensities simply cannot be captured by a linear function. In such settings, one can rely on a more general assumption, namely that spatially corresponding intensities have a *fixed* and *probabilistic* relationship. In other words, one presumes the existence of a joint distribution, $p(I_f, I_m)$, for spatially corresponding image intensities that is location independent and that exposes some probabilistic dependency between intensities (i.e. $p(I_f, I_m) \neq p(I_f)p(I_m)$). Under such assumptions, the registration task can be directly associated [130] to the problem of maximizing the mutual information (MI) between corresponding image values,

$$S_g^{MI}(I_f, I_m^\downarrow) = H(I_f) + H(I_m^\downarrow) - H(I_f, I_m^\downarrow) \quad (2.20)$$

where $H(I)$ is the *marginal* entropy of a probability density function, $p(I)$, and $H(I_f, I_m^\downarrow)$ is the *joint* entropy of the probability density function, $p(I_f, I_m^\downarrow)$.

The normalized variant of MI is referred to as Normalized Mutual Information (NMI)

as is commonly expressed as,

$$S_g^{NMI}(I_f, I_m^\perp) = \frac{H(I_f) + H(I_m^\perp)}{H(I_f, I_m^\perp)}. \quad (2.21)$$

There has been a vast amount of work [30, 35, 78, 87, 97, 137, 155, 160, 161, 167] involving the use of MI in medical image registration contexts. Part of that research effort has been devoted to the accurate and efficient computation of MI, with particular relevance to the technicalities of accurately estimating the joint density function, $p(I_f, I_m)$, while maintaining a reasonable computational complexity. For example, in [171], the authors adopt a Parzen windows strategy for estimating $p(I_f, I_m)$ and reduce computational complexity by relying on a stochastic gradient descent optimization strategy. Stochastic optimization is performed by randomly selecting a small subset of voxel locations at each iteration of the optimization. The corresponding entropies are then estimated with a different random subset of voxels.

Most of the initial works [35, 167, 171] based on MI were designed to address the multi-modal rigid registration of brain volumes. Since most of these modalities are essentially tissue-based (i.e. the intensity of a given voxel is largely dependent on the tissue-types found within the voxel's volume) the use of MI provided a well-justified metric. There were important subsequent efforts [137, 155] in addressing the non-rigid registration of mono-modal images with the use of MI. For example, the work in [137] proposed an MI-based technique for the non-rigid registration of intra-patient mammograms.

More recently, there have been some attempts at employing MI in more challenging multi-modal registration contexts such as the registration of pre-operative MRI to intra-operative US [78, 87], where the authors typically rely on some pre-processing stage for improved robustness. Nonetheless, the use of MI in such approaches is somewhat harder to justify in contexts like these where the image formation models are very different. In particular, the image model of US is *not* exclusively tissue-based (in the sense that a voxel's intensity is *not* largely a function of the tissue-type found within the voxel's volume) and involves many complex artefacts, such as a highly non-homogeneous intensity response.

Incorporating Location Dependency to Global Similarity Functions

Even though similarity functions based on CC or MI can address more general registration contexts where spatially corresponding image intensities are related linearly or in a probabilistic fashion, they remain sensitive to image formation models that vary throughout the spatial domain. Consider the case of the image formation model of an MRI machine involving a *bias field* that corrupts images as a function of the inhomogeneities in its magnetic field. Fig. 2.4 illustrates a MRI brain volume affected by such a bias field, where one can note how the image intensity is generally brighter in the bottom image region. The reason why a non-homogeneous image formation model affects those metrics lies in the fact that both metrics make a *global* assessment of image similarity. In particular, both NCC and MI depend on the global statistics of the image intensities, and have no mechanism for verifying if such statistics do in fact remain constant throughout the spatial domain. A strategy that has been strongly pursued in recent years to address this limitation has been to incorporate the notion of locality in such metrics.

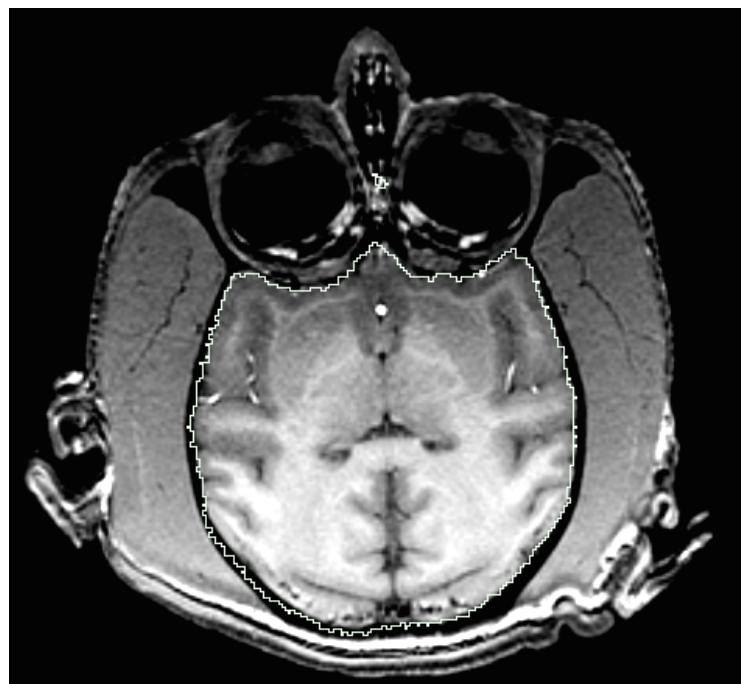


Fig. 2.4 MRI vervet brain volume affected by a bias field. Image obtained from <http://www.slicer.org/slicerWiki/>

One of the most well-known instances of this principle can be found in the registration

algorithm known as the Symmetric Image Normalization Method (SyN) [10], part of the Advanced Normalization Tools (ANTs), where the authors rely on a regional normalized cross correlation (RNCC) similarity function for resolving a diffeomorphic transformation. The similarity function can be expressed,

$$S_g^{RNCC}(I_f, I_m^\downarrow) = \sum_{\mathbf{x} \in \Omega_z} \frac{(I_f(\mathbf{x}) - \bar{I}_f(\mathbf{x}))(I_m^\downarrow(\mathbf{x}) - \bar{I}_m^\downarrow(\mathbf{x}))}{\sigma_{I_f}(\mathbf{x})\sigma_{I_m^\downarrow}(\mathbf{x})} \quad (2.22)$$

$$\bar{I}(\mathbf{x}) = \frac{1}{|\mathcal{N}_{\mathbf{x}}|} \sum_{\mathbf{x} \in \mathcal{N}_{\mathbf{x}}} I(\mathbf{x}) \quad (2.23)$$

$$\sigma_I^2(\mathbf{x}) = \frac{1}{|\mathcal{N}_{\mathbf{x}}|} \sum_{\mathbf{x} \in \mathcal{N}_{\mathbf{x}}} (I(\mathbf{x}) - \bar{I}(\mathbf{x}))^2 \quad (2.24)$$

where $\mathcal{N}_{\mathbf{x}}$ is the neighbourhood of voxel locations centred at \mathbf{x} .

Similarly, in [138], the authors apply the same principle to MI, where they propose to divide the spatial domain into a fixed set of disjoint neighbourhoods and then evaluate the global value of MI as the average of the neighbourhood specific MI values. Alternatively, in [94], the authors propose to improve the robustness of MI by extending the probability feature space with a random variable representing the spatial location and then compute the MI between image intensities conditioned to the spatial location. Broadly speaking, the metric (referred to as Conditional Mutual Information) can be expressed as the weighted average of MI obtained in spatially constrained regions, where the weights roughly correspond to the number of pixels found in the region. The metric is burdened by an increase in computational complexity. However, it provides an improvement in robustness to non-homogeneous image formation models and to contexts where objects with a common intensity response appear in different parts of the image.

In [80], the authors pursued the question of how to express the analytical limit of MI as the spatial region of interest approaches the voxel size. It is clear that one cannot rely on a single pixel pair sample for estimating a probability density. Hence, they locally characterize the image with a first-order Taylor expansion and find that MI (referred to as Local Mutual Information (LMI)) can be expressed as a monotonically decreasing function of the inner angle, $\theta = \angle(\nabla I_f, \nabla I_m^\downarrow)$, of corresponding image gradient orientations,

$$S_g^{LMI}(I_f, I_m^\downarrow) = \frac{1}{|\Omega_z|} \sum_{\mathbf{x} \in \Omega_z} C_d - \log_2 |\sin(\theta(\mathbf{x}))| \quad (2.25)$$

where C_d is a constant that depends on the dimension of the image.

Such work provides a theoretical justification to go beyond the assumption of intensity correspondence and rather focus on finding a correspondence between features based on first-order image differentials (in this case, the feature is the gradient orientation). The following section describes approaches that *explicitly* rely on a feature-extraction function and then define the similarity function based on dense image features defined in entire image domain.

Dense Image Features

Various image registration algorithms rely on a feature extraction function, ψ , which maps the image values in given neighbourhood to one or more features which can be compared to another feature vector with a pre-defined similarity function. The justification behind feature-based registrations techniques is commonly related to improved robustness with regards to image artefacts and multi-modal settings. The dense features that have been proposed for medical image registration contexts are typically based on the statistical distribution of intensities in a local region, first and second-order derivatives or shape-based descriptors.

For example, image gradients have been used by various registration techniques [3, 65, 68, 119, 131, 145, 146] as dense features of interest. Some works [65] have proposed to assess image similarity based exclusively on image gradients. Other techniques [68, 119, 131, 145] use image gradients in a complementary fashion with an intensity-based metric. For example, in [119] the authors complement the NMI metric with a gradient-based term that is multiplied with the NMI value. The gradient-based term is evaluated as a function of both gradient orientation alignment and corresponding gradient magnitudes. Alternative approaches [3, 146, 156] have proposed evaluating MI over a set of gradient-based image features, as opposed to image intensities.

The idea of evaluating MI over an extended set of features has also been pursued in [156] where the authors evaluate not only the voxel's intensity value, but also a set of scalar values that reflect structural information and are computed based on the first and second-order derivatives of the images. The feature images are then evaluated with an estimation of MI known as α -MI [71] that provides a computationally efficient technique for evaluating MI in a large feature space.

Some recent approaches [69, 127] have recently adopted local self-similarity [148] descriptors as robust local image features for complex multi-modal settings. A local self-similarity descriptor is a feature-vector that evaluates the intensity difference of neighbouring voxels with relation to the intensity of the voxel of interest. For example, in [69] the authors propose a Non-Local Shape Descriptor (NLSD) based on the self-similarity of neighbouring patches. Each component of the NLSD feature vector is evaluated as the Gaussian function of the intensity difference between all pixels found in the patch centred at the voxel of interest and a corresponding neighbouring patch. The authors then propose a local similarity function defined as the NCC of the components in the NLSD feature vector.

Dense feature based registration strategies are typically proposed so as to provide improvements in registration robustness. However, in many cases, this leads to a significant increase in computational complexity, which is related to the extraction of the features from transformed images. In particular, since the extracted image features are functions of neighbouring image intensities, one cannot simply apply a transformation function to the image features as if they were conventional scalar images, but must instead apply the transformation function to the image intensities and then re-compute the image features. In other words, a feature based registration strategy will typically involve some computational overhead related to re-computing the image features with every candidate transformation being evaluated.

2.1.7 Optimization

The choice of optimization strategy is typically based on a trade-off between robustness, accuracy and computational complexity. In the context of medical image registration, there is a dominant trend to address the optimization task with a gradient-descent related strategy. However, there has been significant work in employing direct search methods (e.g. simplex method) when either the computational cost of the evaluating the energy function is minimal and/or the space of transformation parameters is relatively small. Furthermore, there have also been some efforts in employing second-order optimization methods (e.g. Newton's method) for reducing the number of iterations necessary to reach the maxima. For a more detailed analysis on the performance of different optimization strategies in specific image registration context, the reader may refer to [20]. This section reviews the major categories of optimization techniques.

Direct Search Methods

Direct search methods are optimization strategies that rely solely on evaluations of the objective function and do not rely on its derivatives. They are typically based on stochastic principles (e.g. simulated annealing) or graph-based frameworks (e.g. iterated conditional modes).

Stochastic based direct search methods can be found in various image registration strategies, but can be considered better suited for contexts where the transformation parameters are of reduced size (e.g. rigid transformation). For registration contexts involving a large set of transformation parameters, these strategies tend to lead to an impractical computational complexity.

First-Order Methods

First-order optimization methods rely on the first-order differential of the objective function. They represent a wide range of techniques that are commonly referred to as *gradient descent* strategies. Most of these techniques can be expressed at a given optimization iteration k as,

$$\mathbf{W}_T^k \leftarrow \mathbf{W}_T^{k-1} + \lambda^k \cdot \nabla_{\mathbf{W}} E_g(I_f, I_m, \mathbf{T}) \quad (2.26)$$

where \mathbf{W}_T corresponds to the vector of parameters that characterize the transformation function \mathbf{T} , $\nabla_{\mathbf{W}} E_g(I_f, I_m, \mathbf{T})$ is the gradient of E_g with respect to \mathbf{W}_T , and λ defines the *step size* at each iteration.

First-order methods are *greedy* and can only find a *local extrema*. Hence, they generally rely on the assumption that the initial location is within the vicinity of the *global extrema*. In the context of image registration, special care is often devoted to ensuring that the gradient descent strategy does not fall into a local extrema. A common strategy is to use a multi-resolution image pyramid, which is described in Section 2.1.8.

Second-Order Methods

Second-order optimization methods rely on the second-order differential of the objective function. Many of these techniques can be associated to Newton's method. In general, Newton's method is an approach for finding the location of a function's zeros (i.e. $f(x) = 0$). In the context of image registration, it is an iterative technique, somewhat similar to

gradient descent, which finds local extrema defined as the location a objective function's derivative zeros. An iteration k of Newton's method can be expressed as,

$$\mathbf{W}_{\mathbf{T}}^{k+1} \leftarrow \mathbf{W}_{\mathbf{T}}^k + \lambda_n \cdot \mathbf{H}^{-1}(E_g) \cdot \nabla_{\mathbf{w}}(E_g(I_f, I_m, \mathbf{T})) \quad (2.27)$$

where $\mathbf{H}(E_g)$ is the Hessian matrix of the energy function, E_g .

Newton's method may provide an improved rate of convergence with relation to a gradient decent strategy. However, it relies on the assumption that the objective function can be approximated by a *quadratic* function in the vicinity of the extrema. Note that the evaluation of the Hessian matrix is typically more computationally expensive than to the evaluation of the gradient. Thus, even if the method yields an improved rate of convergence, it may still do so at an increase computational cost with relation to a gradient descent strategy. The reader may refer to [20] for an exhaustive comparison of first-order and second-order optimization methods in the context of image registration.

Various second-order optimization methods that can be characterized as an approximation to Newton's method. Specifically, such methods estimate the value of $\mathbf{H}^{-1}(E_g)$ with a computationally efficient expression that continues to provide improved rates of convergence. For example, the registration technique proposed in [116], commonly referred to as the “Demons’ algorithm” and known for its reduced processing times, has been subsequently described [116] as an estimation of Newton's method.

2.1.8 Improvements in Computational Efficiency

Multi-Resolution

One of the most commonly used techniques for improving computationally efficiency and robustness against local extrema in image registration techniques is the use of a multi-resolution image pyramid [90]. The typical setup in a multi-resolution registration approach can be described as N sequential registration stages where each stage involves an image pair with a specific resolution. More specifically, the initial registration stage consists of a coarse image resolution and its solution is used to initialize a subsequent stage involving a finer image resolution. The final registration stage is commonly defined as the original resolution of the input images.

Fig. 2.5 illustrates a conventional multi-resolution image pyramid where higher pyramid

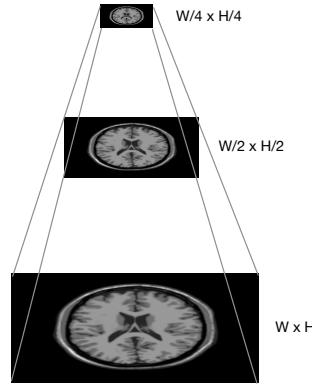


Fig. 2.5 Multi-resolution image pyramid.

levels corresponds to coarser registration stages and are characterized as upsampled variants of the original input image. Multi-resolution image pyramids are heavily based on scale-space theory [90] in which the notion of scale is characterized by the convolution of the image with a Gaussian function of a given standard deviation. Hence, most multi-resolution image pyramids rely on a Gaussian blurring operation of different scale at each registration stage which is then downsampled to the specified resolution.

Voxel Selection

A commonly used technique [15, 18, 21, 49, 57, 99, 112, 114, 123, 139, 167] for reducing the computational complexity of a registration is to prudently select a reduced subset of voxels and restrict the evaluation of the similarity metric to such voxels.

Some of the early works [99, 167] in medical image registration already include this idea of relying on fully randomized voxel selection scheme. In such works, there was no attempt in giving a different priority to each voxel, so as to choose the most informative voxels. Note that the strategy is a very simple approach for reducing computational complexity and has no inherent bias regarding the voxels selected. Nonetheless, a fully random selection of voxels typically leads to a significant loss in registration accuracy. In order to mitigate the penalty in registration accuracy, subsequent methods [21, 49, 57, 114, 139] have developed heuristics and theoretical frameworks for defining the *utility* of a given voxel in resolving the transformation.

Many of the recently proposed voxel selection strategies for intensity-based similarity

metrics lead to a preference for high gradient magnitude locations. The reasoning behind such preference typically highlight that such locations can be associated to increasing variations of the similarity metric and are thus effectively driving most of the image registration. In contrast, the image similarity in low gradient magnitude location will not be greatly affected if the image is slightly displaced. For example, in [139], the authors propose a gradient magnitude subsampling (GMS) approach, where the moving image is subsampled on off-grid locations with a probability of being sampled that is proportional to the gradient magnitude.

Voxel selection strategies are generally proposed so as to provide a similar registration accuracy as conventional full sampling registration approaches with a significantly reduction in computational complexity. However, the advantages of voxel selection strategies have generally been limited to registration contexts with linear transformations. Registration contexts involving complex non-linear transformations (e.g. B-spline transformations) typically suffer a sharper degradation in registration accuracy as the sampling rate is decreased. It also important to note that voxel selection strategies are typically coupled with a “conventional” intensity-based similarity metric (e.g. SSD, NCC, MI) and are not meant to improve the robustness against challenging imaging artefacts, such as non-homogeneous image formation models, occlusion and complex noise models.

2.2 Medical Applications

Image registration techniques can be found in a wide variety of medical image analysis applications, each with a unique set of challenges. Some those applications include the fusion of anatomical images (e.g. CT, MRI) and functional images (e.g. PET, fMRI) [35, 57, 171, 172], segmentation of anatomical regions [36, 38], construction of reference atlases [12, 56, 59, 79, 173, 175], longitudinal analysis of anatomical structures of interest [39, 46, 110], computer assisted interventions [34, 47, 78, 88], among many others. The reader is invited to refer to [54, 62, 72, 98, 113, 118, 150] for exhaustive surveys characterizing the field. This section includes a brief summary of a set of key medical applications of image registration techniques that relate to the work presented in this thesis.

2.2.1 Multi-Modal Fusion of Anatomical Images

One of the most common uses of medical image registration is the rigid alignment of anatomical images, such as CT or MRI volumes, so that spatially corresponding voxels correspond to a common anatomical location. The registered images can then be visually “fused”, allowing the user to view multiple modalities in a common coordinate frame.

A major assumption behind such registration contexts is that anatomical structures do not undergo any non-rigid deformations between the time points where each image was obtained. It is important to note that one also assumes that the image formation model itself does not induce a deformation that is different for each acquisition. Such an assumption is not fully satisfied by various medical imaging modalities. For example, MRI images typically suffer a (small) non-linear deformation induced by the inhomogeneities of the magnetic field [27].

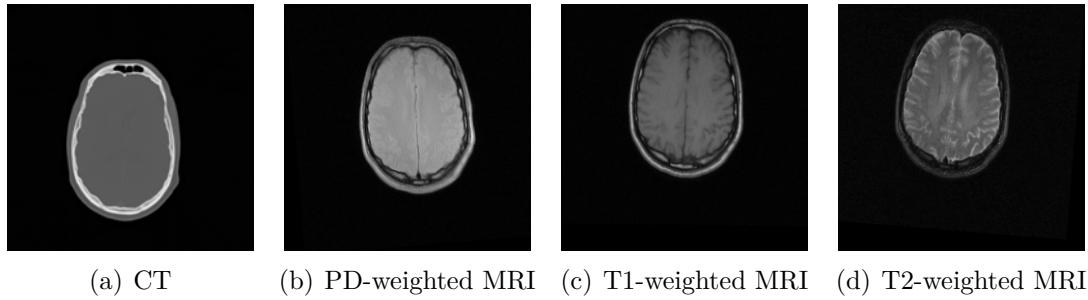


Fig. 2.6 Case involving the rigid registration of brain CT and MRI volumes. Images obtained from the RIRE dataset <http://www.insight-journal.org/rire/>.

An iconic example of multi-modal fusion can be found in the rigid alignment of multi-modal brain volumes. This context has been thoroughly explored by many different techniques [35, 45, 57, 119, 120, 171] and there are publicly available datasets [172] that can be considered a well-accepted validation standard for comparing techniques. Fig. 2.6 illustrates a clinical case involving the rigid registration between a brain CT volume and a set of brain MRI volumes. A majority of the techniques previously proposed for this context [35, 57, 171] have relied on the maximization of NMI, which constitutes an intensity-based multi-modal similarity metric.

2.2.2 Anatomical Segmentation

Image registration techniques commonly play a key role in image analysis pipelines with the objective of segmenting anatomical structures. The basic principle of such segmentation strategies is to rely on a pre-defined expert labelled atlas (template) image where the anatomical region of interest has been manually identified. Some important variants to this approach are discussed in this section.

Some techniques rely on a hard mapping of labels based on the spatially mapped atlas labels. For example, in [36, 38], the authors propose a segmentation algorithm that registers a patient’s brain MRI to a standard brain template, and then resolves the patient’s brain tissue labels via the spatial mapping to the labels in the template obtained from a patient to template non-linear registration.

Other segmentation techniques rely on a probabilistic atlas [56, 100, 147] where each voxel involves a vector of tissue type probabilities. For example, the ICBM 2009 brain atlas [56] includes for each voxel the probability of it being classified as grey matter, white matter or cerebrospinal fluid. Different works [7, 11] have proposed probabilistic segmentation strategies based on the Expectation Maximization (EM) algorithm using a probabilistic brain atlas as a prior probability for the healthy tissue type segmentation of brains.

There are also probabilistic segmentation techniques [51, 81, 158] for the purposes of brain lesion detection that also rely on healthy tissue type probabilities obtained from a registration to a probabilistic brain atlas. In this context, the lesion detection methods rely on the prior knowledge regarding the healthy tissue type probability of each voxel for characterizing a probabilistic model that, in conjunction with multiple image features, allows for the detection of pathologies.

An alternative approach for segmenting anatomical structures of interest involves registering the subject’s image to *multiple* atlases that include a valid segmentation. Such methods [1, 75, 140, 170] are referred to as Multi-Atlas Label Fusion (MALF) techniques. The patient’s segmentation labels are obtained by warping the labels in the pre-segmented atlases to the patient’s space. One must also define a fusion scheme which describes how multiple template labels are jointly used to resolve the patient’s segmentation label. For example, a commonly used fusion scheme is to adopt the label value that was obtained from a majority of the reference templates.

2.2.3 Longitudinal Analysis

Non-rigid registration techniques can be commonly found in intra-patient longitudinal studies [8, 9, 39, 126, 137, 174], where a given subject is imaged at multiple time points and the registration is used to infer the physical deformation the subject underwent across time. This strategy has been pursued with various image modalities and anatomical structures of interest. For example, in [137] the authors proposed a non-rigid registration technique for inferring the deformation between MR mammograms. In [39], the authors propose a technique for inferring the deformation between lung CT volumes at exhale and inhale states. In [8], the authors rely on a symmetric diffeomorphic registration technique for inferring the amount of grey matter atrophy suffered by a patient across imaging time points.

2.2.4 Image Guided Interventions

Computer assisted interventions are contexts where image registration can play a key role in multiple stages. Most computed assisted clinical procedures are characterized by a pre-operative imaging modality followed by an intra-operative imaging acquisition which may be used for either tracking the location of a region of interest, characterizing physical deformation undergone after the pre-operative image was acquired and/or exposing key anatomical regions that can be solely exposed by an intra-operative modality.

The key responsibility of any image registration involved in such procedures is to resolve the spatial mapping between two or more imaging acquisitions. At an initial planning stage, prior to the clinical intervention, multiple image modalities (e.g. T1-weighted MRI, CT, DTI, PET, fMRI, etc.) may be used for diagnostics and planning of the procedure. To allow the clinician to interpret all pre-operative modalities in a common coordinate space, one must rely on a multi-modal rigid registration strategy. Once the clinical intervention is ongoing, one must rely on an image registration technique to rigidly or non-rigidly bring the pre-operative imaging data in spatial correspondence with the intra-operative data. The clinician may rely on such intra-operative registration to either: 1) allow the interpretation of all pre and intra-operative modalities in a common coordinate space, or 2) characterize the physical deformation that has taken place since the pre-operative images were acquired.

Image Guided Radiation Therapy

Image Guided Radiation Therapy (IGRT) [50, 76, 83, 115] also known as Image Guided Radiosurgery (IGRS), is an alternative to standard tumour resection surgery and involves the use of accurately targeted radiation beams to treat tumours. In order to account for patient displacement throughout the radiation treatment, radiosurgery systems typically employ an image-guided solution that continuously tracks the region of interest and allows a robotic arm to continuously reposition the radiation beam. For example, the CyberKnife™ System[28] is a commercial solution that provides continuous skull tracking with the use of X-Ray imaging for the targeting of intra-cranial tumours. The tracking of the skull, also known as motion correction, is driven by the registration between a pre-intervention CT acquisition and intra-intervention X-Ray images. The registration resolves the transformation between corresponding coordinate spaces and leads to an update of the robotic arm location.

Image Guided Neurosurgery

One of the motivating clinical contexts for this work can be found in Image Guided Neurosurgery (IGNS), where the physician relies on pre-operative imaging and a tracked pointer (whose location is known and displayed onto the pre-operative volume) for guidance throughout the surgical procedure. Unfortunately, brain movements during open-skull operations are known to reduce the utility of image guidance. Intra-operative imaging modalities have thus been proposed for improved guidance accuracy. Some modalities, such as intra-operative MRI, involve a prohibitive cost, as well as requiring major modifications to the operating room and surgical procedure. Alternatively, intra-operative ultrasound (iUS) provides a highly appealing option given its ease of use and relatively inexpensive costs.

Multi-modal registration plays a fundamental role in this context. In particular, one can leverage a registration technique for resolving the deformation between a pre-operative MRI volume and an iUS so as to update all pre-operative information and compensate for soft tissue deformations found through the operation. This registration task is particularly challenging given the widely different nature of the modalities involves, and there has been a significant amount of work [47, 78, 127, 128] devoted to it. As such context constitutes a critical component of this work, it will be further detailed in Chapter 4 prior to presenting corresponding experimental results.

2.3 Validation of Registration Techniques

Validation of medical image registration algorithms is an ongoing challenge faced by the research community which arises from an absence of ground truth in real medical contexts. There have been many different approaches proposed for estimating the geometric accuracy of registration algorithms and the topic has been discussed in detail in some works [109, 132]. One should note that even though the geometric error is the direct and natural measure of accuracy of a registration task, its relevance is closely related to the clinical task at hand. For example, some clinical tasks may indeed require an accurate spatial mapping of anatomical regions found throughout the entire image domain, and a measure of geometric error evaluated across the domain may indeed be a suitable validation strategy. Other tasks may instead focus on tracking a particular region of interest, and choosing a performance metric evaluated over the entire domain may poorly reflect the accuracy in the region of interest. This section describes the main validation strategies employed in medical image registration contexts.

2.3.1 Simulated Registration Tasks

Some validation strategies [14, 18, 144, 169] rely on synthetic or simulated medical image volumes generated with a known ground-truth spatial deformation. Hence, they allow for a direct computation of the geometric error at any location in the spatial domain. Unfortunately, simulated medical images recreate their real counterparts with a limited accuracy and typically do not incorporate all the complex artefacts and imperfections found in the true image formation model. Additionally, in the context of non-rigid deformations, one is faced with the challenge of defining a physically plausible transformation that reflects the biomechanical properties of the anatomical structures of interest. For example, in [144], the authors propose a finite-element method for simulating gold standard deformations so as to validate a breast MRI non-rigid registration technique.

2.3.2 Distance between Homologous landmarks

An alternative strategy is one that relies on true medical volumes found in real clinical cases. However, one is rapidly burdened by the task of identifying a suitable gold standard. The task of defining a valid ground truth (i.e. the “true” transformation function) is

rendered practically impossible in the context of complex non-linear transformations with the extreme scenario where a displacement must be specified at each voxel location.

In the absence of a ground-truth transformation function, some authors [47, 99, 109, 117, 119, 128] rely on manually selected homologous landmarks and evaluate the mean distance (or any other related statistics) as the registration performance metric. Furthermore, in contexts that involve simple transformations (e.g. rigid transformation), one can also make use of such point correspondences for characterizing the geometric error throughout the spatial domain [55].

It is important to note that the manual selection of homologous landmarks can be strongly biased by each expert and the prior knowledge he/she relies on for identifying the landmarks. In particular, a clinical expert may rely on anatomical knowledge regarding a specific anatomical region of interest for assessing its location, even though there might be no visual information illustrating a correspondence. Furthermore, if the expert has complete freedom to choose landmarks he/she might do so with an arbitrary spatial distribution that does not properly capture registration accuracy throughout the spatial domain.

In order to reduce the effect of expert bias and poor spatial distribution of landmarks, validation strategies may enforce constraints on the landmarks that can be selected. For example, in [109] the authors propose a state-of-the-art semi-automatic strategy that generates a well-distributed large set of homologous landmarks and relies on minimal user interaction for correctness.

2.3.3 Agreement between Spatially Mapped Labels

In the absence of a complex ground truth deformation field and with the goal of bypassing any manual identification of point correspondences, there has been significant work [2, 43, 84, 89, 94, 134, 141, 166] in validating registration techniques based on the overlap of expert segmented anatomical regions and tissue types. Such methods typically involve a pre-existing dataset of expert-annotated medical volumes. The approach has received some criticism [132] which underlines that, in some contexts, an excellent agreement between anatomical labels can be brought forward by an implausible transformation function. Nonetheless, for contexts where the anatomical labels correspond to unique and localized anatomical regions, as opposed to general tissue types, the validation approach can provide an informative surrogate measure with regards to geometric accuracy. For example, in [84],

the authors evaluate the performance of fourteen state-of-the-art non-linear registration techniques by relying on four different datasets of T1-weighted MRI brain volumes with corresponding expert segmentations of various anatomical regions (e.g. cortical foldings). Fig. 2.7 illustrates the four different datasets of expert-labelled MRI brain volumes used. Registration performance is evaluated by first registering a patient’s MRI volume to a second patient’s MRI volume, and then evaluating the agreement between spatially mapped anatomical labels.

2.4 Summary

Image registration is a problem that extends to a wide variety of applications and involves decades of research efforts. The diversity of challenges faced in each particular context has lead to large family of techniques that aim to improve each of the registration components including the energy function to be optimized, the choice of image features, the transformation model used to deform images, the image interpolation technique and the optimization strategy used. The overview of registration methods presented in this chapter illustrates the particular importance of the similarity metric and transformation model used, and how these two critical components largely characterize the type of problems that can be addressed.

As automated medical image analysis continues to develop and become a more critical part of standard clinical procedure, registration of medical volumes stands out as a critically important component that is present in a vast majority of such applications. The overview of medical image registration contexts presented in this chapter listed some key examples and the challenges that may be encountered. This chapter also illustrated the major difficulties in defining a validation strategy that is suitable for a particular medical context, and how one is divided between a synthetically generated yet highly accurate strategy or the use of a dataset of real clinical cases with manually generated ground truth.

An important phenomenon that stands out in the overview of previous work is the conflict between robustness and computational efficiency. In particular, there seems to be a common trend towards improving registration robustness by increasing the dimensionality of the image feature space, which unfortunately also leads to a natural increase in computational complexity and processing time. This provides a fundamental challenge for medical applications with time sensitive constraints such as the ones found in image guided

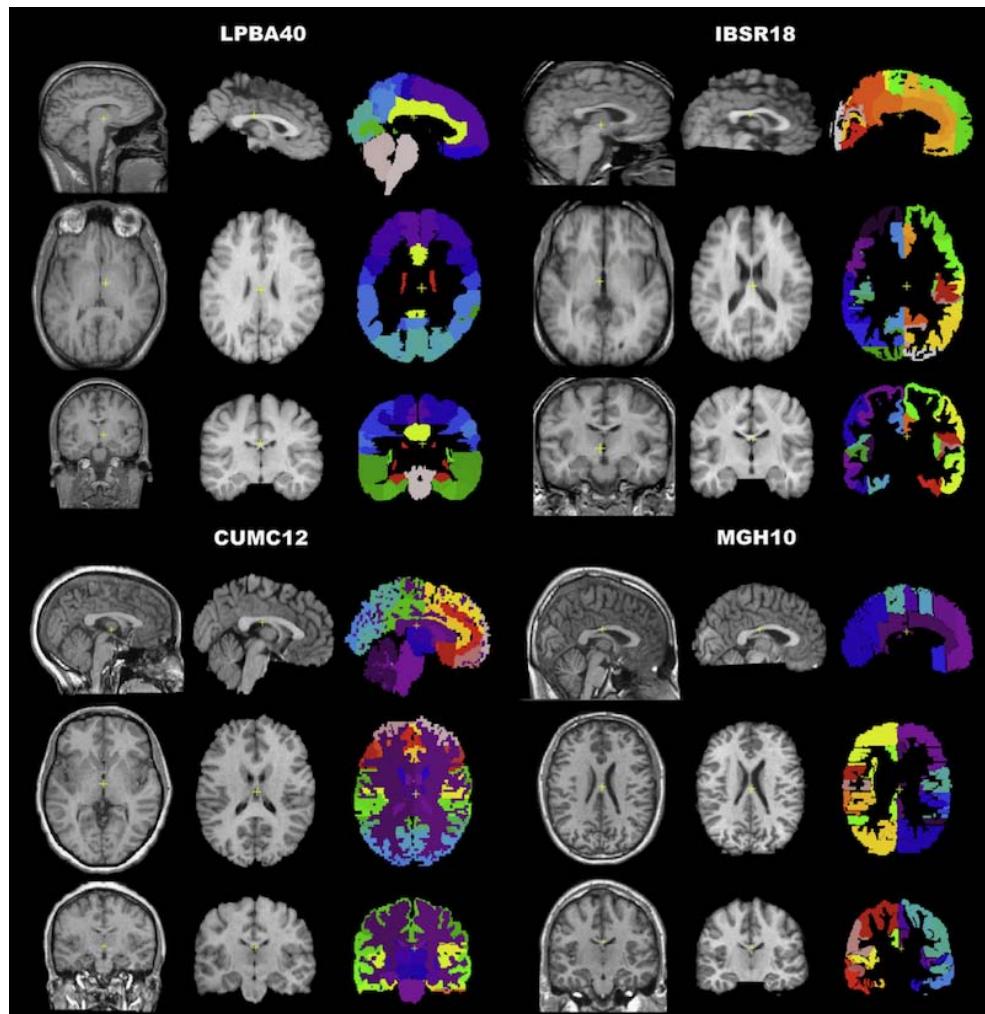


Fig. 2.7 Publicly available datasets of expert labelled brain MRI volumes used for the evaluation of non-linear registration techniques performed in [84]. Each quadrant illustrates the MRI volumes and anatomical labels from a specific dataset. The left column of each quadrant illustrates a T1-weighted head MRI. The middle column of each quadrant illustrates the same T1-weighted MRI after removal of non-brain tissue by brain masking. The right column of each quadrant illustrates the anatomical labels identified in the T1-weighted MRI. Image obtained from <http://www.mindboggle.info/data.html>.

interventions where real time or near real time performance is critically important for a successful integration with a clinical procedure.

The work presented in this thesis seeks to achieve the same type of advantage in registration robustness brought forward by feature-based techniques, but avoiding a penalty

in terms of computational complexity. Instead, it seeks a *reduction* in processing times by relying on a sparse set of data, similar to pixel selection strategies used with intensity-based metrics. The basic principle adopted for achieving such goals is to rely on image gradient orientations at prudently selected locations as a sparse and compact representation of the anatomical structures of interest in an image. The following chapter provides a detailed description of two proposed techniques that follow such principle and following chapters provide a series of quantitative results that illustrate the advantages brought forward with relation to competing state-of-the-art methods.

Chapter 3

Image Registration based on Gradient Orientations of Minimal Uncertainty

The previous chapter presented a general overview of various image registration techniques proposed for medical contexts. An important trend that was identified in such chapter was the use of multi-dimensional image features, proposed so as to achieve improved registration robustness. Unfortunately, the use of a large number of image features was shown to also involve a significant increase in computational complexity corresponding to the cost of computing the image features of a transformed image for every candidate transformation function. Needless to say, there are many time-sensitive clinical applications which require registration solutions with significantly reduced computationally complexity *in conjunction* with strong robustness against complex image models.

This chapter presents the theory and methodology regarding two closely related registration techniques that are motivated by such challenging time-sensitive applications and which largely rely on image gradient orientations as features for assessing the quality of an image match. The basic principle of using image gradients as a driving feature for image registration is not, in and of itself, a novel idea. However, such efforts [65, 68, 80, 101, 119] have generally been limited to a *dense* evaluation of the similarity metric throughout the image domain, and there has also been no effort in minimizing the computational overhead introduced when evaluating the gradient orientations of a transformed image. The

framework presented in this thesis directly addresses these issues by focusing on a sparse set of locations that relate to structures of interest and illustrating how the evaluation of transformed image gradients can be performed with computational efficiency.

There are two key intuitions that are behind the registration techniques developed in this chapter. The first is that a majority of the image domain is characterized by visually homogeneous regions involving locations that are virtually indistinguishable from each other, and it is thus more informative and profitable to focus on visually non-homogeneous regions. The second is the principle of characterizing such regions of interest with a primitive *geometric* descriptor inferred from the image, instead of relying on the region's image intensity profile. The work in this thesis adopts what is perhaps the simplest geometric descriptor that can be inferred from an image region, that is, the orientation/direction in which change is occurring in the region.

In this work, the principle of focusing on image regions where “something is happening” is directly associated to accurately identifying locations corresponding to an interface between homogeneous regions (i.e. boundaries). Such task is considered essentially equivalent to the problem of edge detection, where the goal is to identify image boundaries that correspond to boundaries in the underlying scene. Nonetheless, the problem of accurately and robustly identifying the location of anatomical boundaries represents a non-trivial challenge by itself, with a significant amount of research devoted to it. For the purposes of a gradient orientation based registration and where inferred boundaries are used for restricting the locations where a similarity metric will be evaluated, the sensitivity to edge detection performance is shaped by the complexity (i.e. degrees of freedom) of the transformation model being resolved. For example, a *global*¹ parametric transformation model (e.g. rigid transformation) involves a small set of parameters and can be analytically solved with a few point correspondences. In other words, one can solve a rigid transformation model with a few geometric constraints. Furthermore, if the geometric constraints are free of errors, then the resulting transformation will also have no error and there will be no benefit in increasing the number of geometric constraints. Alternatively, if the geometric constraints involve some error or degradation, an increase in the number of constraints will lead to a reduction in the error of the transformation. The first method in this chapter addresses the rigid registration of multi-modal volumes based on such principles. It characterizes

¹A rigid transformation model is *global* since the scope of its parameters is global and not local. In direct contrast, a B-spline transformation is *local* since its parameters have a strictly local scope.

the confidence on the gradient orientations with a formal analysis on the effect of additive Gaussian noise on the computed orientations. It then relies on such indicator of confidence so as to prudently select a sparse set of locations which will accurately drive the registration process.

In direct contrast to a global parametric transform, a non-linear transformation model (e.g. B-spline transformation) typically involves a large set of locally-scoped parameters and therefore requires an increased number of geometric constraints for it to be resolved. Such transformation models provide some degree of local freedom, where sufficiently far spatial regions have fully decoupled displacements. This may quickly lead to an ambiguity when relying on the constraints evaluated by gradient orientation alignment. Specifically, any ambiguity in the precise location of the boundary will lead to an ambiguity in the local deformation. Unfortunately, most of the images encountered in practice do in fact involve such ambiguity since they expose boundaries with an image gradient that spans over a certain width. In other words, the image gradient by itself does not provide a pixel compact identification of the boundary's location. This is a challenge for a registration framework based on gradient orientation alignment since it constitutes an ambiguity in both the selection of locations for the evaluation of the metric and in the geometric accuracy of the metric with relation to a second image. Such limitations can be directly addressed by adopting an explicit dependency on a pixel compact edge detection method which can accurately define the locations to be evaluated. Additionally, the ambiguity in the similarity metric itself can be restricted by also including a distance-based term which penalizes locations that are far from an inferred boundary location. These are the principles that characterize the second proposed method. In particular, the method involves a diffeomorphic (non-rigid) registration technique which identifies a sparse set of boundary locations in each image and consequently seeks the transformation that jointly maximizes alignment with their spatially corresponding gradient orientations and minimizes the distance to the nearest spatially corresponding boundary. In order to enforce a hard constraint on the degrees of freedom of the non-linear transformation model, the proposed technique also relies on a coarse spatially discretized velocity field. and also relies on symmetric consistency, where the similarity metric is evaluated on selected locations identified in the fixed image domain and in the moving image domain.

In summary, the fundamental motivation behind both techniques lies in exploiting the sparsity of boundary information, as evaluated through gradient orientation alignment, in

a computationally efficient fashion, while relying on geometric consistency constraints for facilitating the optimization tasks. It is clear that the two proposed methods are in reality variants of a common framework. Nonetheless, there are significant technical differences and their methodologies are therefore described in separate sections.

The techniques described in this chapter are proposed with the aim of being generally applicable to multiple mono-modal or multi-modal contexts involving image formation models that can expose anatomical structures or boundaries (as opposed to functional imaging modalities). The subsequent chapters illustrate the generality of the proposed techniques with important contributions in registration performance in the contexts of: the rigid registration of pre-operative MRI to intra-operative US for neurosurgical procedures, the rigid registration of CT and MRI brain volumes, the inter-patient non-rigid registration of MRI brain volumes for automatic anatomical labelling, and the registration of patient to atlas MRI brain volumes for the generation of healthy tissue probability maps.

3.1 Multi-Modal Rigid Registration based on Gradient Orientation of Minimal Uncertainty

This section describes a framework for the use of gradient orientation as features of interests in a rigid registration context. Similar to related work [48, 65, 68, 80, 101, 119], a local similarity function based on gradient orientation alignment is adopted. The local similarity metric is maximal when corresponding gradients are parallel and minimal when they are perpendicular.

In this work, a significant effort is devoted to characterizing the reliability or certainty of such image features across the spatial domain, so as to focus computation to areas of low uncertainty. This constitutes a direct contrast with other techniques that simply evaluate the similarity measure uniformly throughout the entire domain. In particular, the computation of gradient orientations is cast as a Bayesian inference problem afflicted by image noise, and this effort leads to the theoretical result that the uncertainty on the estimated gradient orientation does in fact vary throughout the image. Specifically, the analysis adopts the assumption of additive Gaussian noise on voxel intensities and describes how it leads to an orientation variance that is spatially non-homogeneous. An important outcome of this analysis is that the corresponding gradient magnitude is an indicator of the orientation variance. In particular, the analysis shows that normalizing the gradient

magnitude with the norm of its corresponding derivative kernel allows one to compare the uncertainty across scales, and one can therefore define a common selection criterion for all scales. This results leads to a multi-scale selection scheme where a gradient magnitude threshold (corresponding to an uncertainty threshold) is defined *a priori* at the finest scale and gradient magnitude thresholds of coarser scales are automatically computed so that they correspond to an identical uncertainty threshold. Such a selection scheme permits a robust and accurate coarse to fine registration strategy that quickly drives the optimization to the neighbourhood of the solution while focusing on confident features at each scale.

Previously proposed gradient orientation based metrics [65, 68, 80, 101, 119] have generally been characterized by a fixed local similarity function, which reflects how tightly orientations should be aligned. In particular, various previously proposed metrics [65, 80, 119] evaluate gradient orientation alignment as the inner product of the normalized image gradients, which also corresponds to the cosine of the inner angle between corresponding gradient orientations (i.e. $\cos(\Delta\theta)$). Unfortunately, such a metric does not provide a high discriminability between “somewhat well aligned” and “tightly aligned” gradient orientations, and an inner angle of 14° already yields a metric value corresponding to 97% of the maximum possible value. The similarity metric presented in this work can be parametrized with different matching criteria and quantitative results will demonstrate that imposing a strict matching criterion on rigid registrations can lead to improvements in registration accuracy.

In order to bypass the evaluation of transformed interpolated voxel intensities and corresponding image derivatives, the proposed framework makes use of a computationally efficient evaluation of transformed image gradients. Specifically, the image gradient of the original (undeformed) image is transformed by multiplication with the Jacobian of the transformation function, a linear mapping known as the *pullback* in differential geometry. In contexts involving rigid transformations, the Jacobian is simply the rotation matrix and is thus constant across the spatial domain.

3.1.1 Methodology

This section describes the proposed rigid registration method, referred to as *GOA* (GOrientation Alignment), and its three main components: 1) a local similarity function evaluating the resemblance of gradient orientations, 2) a multi-scale gradient orientation selection strategy that identifies locations of reduced uncertainty in the fixed image domain,

and 3) a computationally efficient technique for estimating gradient orientations of the transformed moving image.

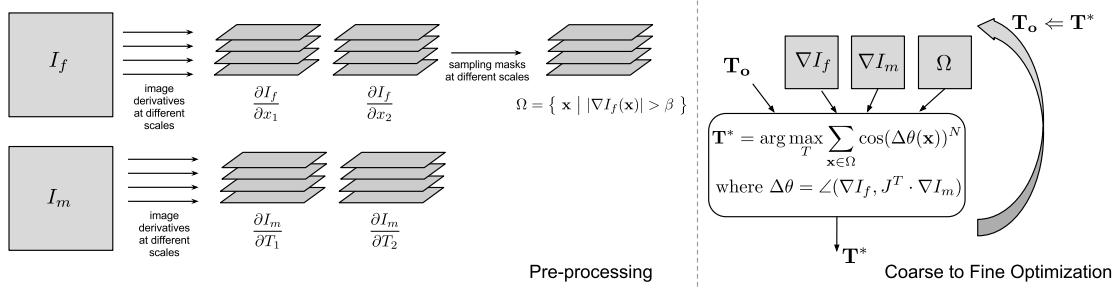


Fig. 3.1 Algorithmic diagram of proposed rigid registration method, referred to as *GOA*. The block diagram illustrates the pre-processing stage and optimization stage of the approach. The pre-processing stage involves the computation of image gradient at various image scales and consequently identifying the sampled locations of interest, Ω , in the fixed image domain and for each image scale. The optimization stage involves the maximization of mean gradient orientation alignment at subsequent image scales and evaluated on the locations of interest. The multi-resolution approach first addresses the coarsest image scale, initializes the next stage with the obtained solution, and proceeds sequentially until reaching the finest image scale.

The algorithm, illustrated in Fig. 3.1, consists of two stages. The first is a pre-processing stage where the image derivatives of both images are computed at multiple scales and where sampling masks (i.e. locations of interests) in the fixed image domain are computed for based on gradient orientation uncertainty and at each image scale. The second stage is a coarse to fine optimization stage in which the local similarity function is evaluated over the selected locations and where each scale is initialized with the transformation obtained in the previous (coarser) scale.

Note the proposed method restricts the voxel selection to the fixed image domain. In that sense, the method can be considered *asymmetric*, and each of the images being registered plays a unique role. This particular choice will be justified in the following chapters, where quantitative benefits in registration performance in multi-modal contexts are presented. It is important to highlight that a *symmetric* variant of the proposed technique is relatively straightforward to develop and may be particularly relevant for mono-modal contexts where both images share a common image formation model. In the following sections, a detailed description of each component as well as of the algorithmic implementation is presented.

Local Similarity Measure

The method relies on a localized similarity function based on gradient orientation alignment that includes a parameter that characterizes its matching criterion. In particular, gradient orientation similarity is evaluated as,

$$s_l(\nabla I_f(\mathbf{x}), \nabla I_m^\downarrow(\mathbf{x}); N) = \cos(\Delta\theta(\mathbf{x}))^N \quad (3.1)$$

$$= \left\langle \frac{\nabla I_f(\mathbf{x})}{|\nabla I_f(\mathbf{x})|}, \frac{\nabla I_m^\downarrow(\mathbf{x})}{|\nabla I_m^\downarrow(\mathbf{x})|} \right\rangle^N \quad (3.2)$$

where $\Delta\theta = \angle(\nabla I_f, \nabla I_m^\downarrow)$ is the inner angle between ∇I_f and ∇I_m^\downarrow ; and $N \in \mathbb{N}$ is the metric's parameter. Notice that the same metric can also be expressed as the inner product between normalized gradients raised to the N -th power. The value of N is an important parameter that relates to the existence of matched boundaries with inverted image intensity transitions, as well as the strictness or selectivity involved in evaluating gradient orientation alignment.

In a general multi-modal context, matched locations may have inverted intensity transitions. For example, a boundary in a first modality may involve a transition from white to black, while its corresponding boundary in a second modality may involve a transition from black to white. In order to reflect such possibility, one can adopt an even value of N , in which gradient orientations with inner angle of π are also rewarded. However, in contexts where it is known *a priori* that matched boundaries must have non-inverted intensity transitions (e.g. mono-modal registration), one can adopt an odd value for N which penalizes corresponding gradients with inverted transitions.

Additionally, the parameter N also characterizes how *strict* or *selective* the matching criterion is. Hence, metrics with large values of N only reward tightly aligned orientations. On the other hand, metrics with small values of N tend to measure the projection of a normalized gradient onto another. In that sense, the parameter N can be considered a tuning parameter of the proposed method. The quantitative results shown in Chapter 5 demonstrate that employing a highly selective metric in the context of rigid registrations can lead to gains in accuracy.

Multi-Scale Selection of Gradient Orientations

Similar to other registration frameworks, the proposed method relies on a multi-resolution strategy where coarse resolution images are registered first and their solution is used to initialize a subsequent (finer) image resolution registration stage. However, the proposed method also relies on a voxel selection strategy which identifies locations of interest where the similarity metric will be evaluated. It is thus of critical value to adopt a *consistent* voxel selection strategy across scales which is based on a common criterion.

This section provides an analysis of gradient orientation degradation brought forward by image noise, which consequently provides an indicator of gradient orientation uncertainty which can be compared across image scales. In particular, the posterior probability of image gradient orientations is derived based on voxel intensities subjected to additive Gaussian noise. This analysis demonstrates that the gradient magnitude is an indicator of the orientation variance (relative to other orientations in the same image) and that normalizing the gradient magnitude with respect to its corresponding derivative kernel also allows one to define a “scale-independent” indicator of gradient orientation uncertainty. The details regarding such derivation are outlined below.

Inferring Gradient Orientations on Images with Noise Consider a discretized 2D image, $I : \mathbb{N}^2 \rightarrow \mathbb{R}$, affected by i.i.d. additive Gaussian noise and whose formation model can be expressed for each pixel as,

$$I[i, j] = F[i, j] + \epsilon[i, j] \quad (3.3)$$

where F is an *undegraded* (i.e. noiseless) image, i, j is a 2D voxel index and $\epsilon[i, j]$ is an i.i.d. Gaussian random variable with variance σ^2 .

The probability density of a pixel intensity is expressed as,

$$p(I[i, j] | F[i, j]) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(I[i, j] - F[i, j])^2}{2\sigma^2}\right) \quad (3.4)$$

Each image derivative, $\partial I / \partial x$ and $\partial I / \partial y$, can be estimated with a linear operator (e.g. central-difference, Gaussian derivative, etc.) that is fully characterized by its corresponding kernel, \mathbf{K}_x and \mathbf{K}_y . Hence, each estimated image derivative, $\hat{\partial}I / \partial x$ and $\hat{\partial}I / \partial y$, resulting from the discrete convolution of its kernel with the image, will follow a normal distribution

(given that the voxel intensity has a normal distribution). In this work, the analysis is restricted to operators that act solely on voxels in the spatial dimension of interest². For example, one may adopt a central-difference operator to estimate the derivative along the x dimension, and such operator only relies on the previous and next voxel along the x axis. For such types of linear operators, the estimated image derivatives are also independent. Finally, if one assumes that the kernels are just rotated variants of each other and thus have the same norm-2 ($\|\mathbf{K}_x\|^2 = \|\mathbf{K}_x\|^2 = \|\mathbf{K}\|^2$), one can express their conditional density as,

$$p(\hat{\nabla}I[i, j] | \hat{\nabla}F[i, j]) = \frac{1}{2\pi(\|\mathbf{K}\|^2\sigma^2)} \exp\left(-\frac{\|\nabla I[i, j] - \nabla F[i, j]\|^2}{2\|\mathbf{K}\|^2\sigma^2}\right) \quad (3.5)$$

where $\hat{\nabla}I = (\hat{\partial}I/\partial x, \hat{\partial}I/\partial y)$ is the computed image gradient and $\hat{\nabla}F$ is the corresponding gradient of the undegraded image.

The joint density of the image gradient magnitude, $m = |\hat{\nabla}I|$, and gradient orientation, $\theta = \angle(\hat{\nabla}I)$, can be estimated through the distribution,

$$p(m, \theta | r, \phi) = \frac{m}{2\pi\|\mathbf{K}\|^2(\sigma^2)} \exp\left(-\frac{m^2 - 2mr \cos(\theta - \phi) + r^2}{2\|\mathbf{K}\|^2\sigma^2}\right) \quad (3.6)$$

where r and ϕ are the magnitude and orientation of $\hat{\nabla}F$ and the voxel index $[i, j]$ is not shown for legibility purposes.

The posterior of ϕ is derived by solving the following integral,

$$\begin{aligned} p(\phi | m, \theta) &= \int_r \frac{p(m, \theta, r, \phi)}{p(m, \theta)} dr \\ &= \int_r \frac{p(m, \theta | r, \phi) \cdot p(r, \phi)}{p(m, \theta)} dr \\ &\propto \int_r p(m, \theta | r, \phi) dr \end{aligned} \quad (3.7)$$

where a uniform prior for (r, ϕ) is adopted (i.e. there is no prior knowledge on gradient magnitude and orientation).

Solving the integral in the previous expression, as well as for the normalization constant,

² Note that some operators (e.g. Sobel) would not satisfy such constraint.

one obtains the resulting expression,

$$p(\phi \mid m, \theta) = \frac{e^{-\frac{-m^2 \sin^2 \Delta}{2||\mathbf{K}||^2 \sigma^2}} \cdot \Phi(\frac{m \cos \Delta}{|\mathbf{K}| \sigma})}{\pi \sqrt{\sigma |\mathbf{K}|} \cdot I_0(\frac{m^2}{4||\mathbf{K}||^2 \sigma^2}) \cdot e^{-\frac{m^2}{||\mathbf{K}||^2 \sigma^2}}} \quad (3.8)$$

where $\Delta = \phi - \theta$ is the difference between the undegraded orientation and the observed orientation, $\Phi()$ is the cumulative distribution function of a Gaussian random variable, and I_0 is the modified Bessel function of the first kind of order zero.

Note that even though the expression found in Equation 3.8 is relatively complex, all non-constant terms can be re-expressed as functions of $\frac{m}{|\mathbf{K}| \sigma}$. Fig. 3.2 illustrates the density with different values of $\frac{m}{|\mathbf{K}| \sigma}$ and shows that it is simply a *unimodal directional density* whose variance is a monotonically decreasing function of $\frac{m}{|\mathbf{K}|}$. In other words, $\frac{m}{|\mathbf{K}|}$ provides a common indicator of the precision (i.e. inverse of variance) of a given gradient orientation across kernels of different scales and without knowledge of the noise variance, σ . Note that without knowledge of σ we are left unable to evaluate absolute probabilities or confidence intervals regarding gradient orientations, and we are also unable to compare the gradient orientation uncertainties between different image modalities.

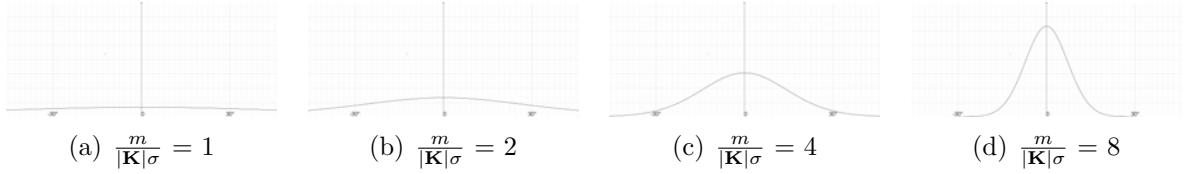


Fig. 3.2 Posterior density of gradient orientation conditioned to different values of $\frac{m}{|\mathbf{K}| \sigma}$.

Multi-Scale Uncertainty Threshold The proposed method relies on the obtained indicator of uncertainty in a multi-scale orientation selection scheme where only locations whose uncertainty is below a given threshold are selected. In particular, a gradient magnitude threshold, β_1 , is defined *a priori* for the finest image scale, and corresponding thresholds for coarser scales are directly obtained. In a multi-scale approach, we have unique kernels for each scale and we can equalize the orientation posterior of different scales. Solving for two scales, k and l , one can derive the following relationship between corresponding

magnitudes,

$$\frac{m_k}{|\mathbf{K}_k|} = \frac{m_l}{|\mathbf{K}_l|}. \quad (3.9)$$

In other words, a gradient orientation, ϕ_k , computed at scale, k , with magnitude, m_k , has the same variance as an orientation ϕ_l , computed at scale l , with magnitude, $m_l = \frac{|\mathbf{K}_l|}{|\mathbf{K}_k|} \cdot m_k$. The threshold value for all scales can therefore be obtained by multiplying β_1 with the ratio of norms of corresponding kernels,

$$\beta_n = \frac{|\mathbf{K}_n|}{|\mathbf{K}_1|} \cdot \beta_1. \quad (3.10)$$

Fig. 3.3 illustrates the multi-scale pixel selection strategy applied to a coronal slice extracted from a synthetic T1-weighted MRI volume obtained from BrainWeb [86]. A four-scale setup is demonstrated, where a gradient magnitude threshold at the finest scale has been defined such that the top 10% pixels with highest gradient magnitude are selected. The thresholds corresponding to coarser scales are automatically obtained by multiplication with the ratio of kernel norms. Fig. 3.3 illustrates the computed gradient orientations and magnitudes, as well as the selected gradient orientations. Notice that the percentage of selected locations increases as the scale of the operator increases. Hence, even if one selects a small percentage of highly confident locations at the finest scale, one can still make use of most of the locations at coarser scales without sacrificing certainty on gradient orientations.

This section described a principled approach for selecting locations of interest across image scales based on a minimal gradient orientation uncertainty, which provides a principle technique for focusing the registration method to regions that more accurately reflect the geometry of the underlying scene. However, the proposed registration method involves the evaluation of gradient orientation alignment on corresponding points, and thus also requires the computation of image gradient orientation of a transformed image. The next section describes the linear mapping, referred to as the pullback differential, involved in evaluating image gradients in a transformed coordinate space and how it can be evaluated in a linear registration context so as to minimize computational complexity.

3.1.2 Transformation of Image Gradients

A straightforward approach for evaluating the gradient orientations of a transformed moving image, $I_m^\downarrow = I_m \circ \mathbf{T}$, involves first evaluating the pixel intensities by some interpolation

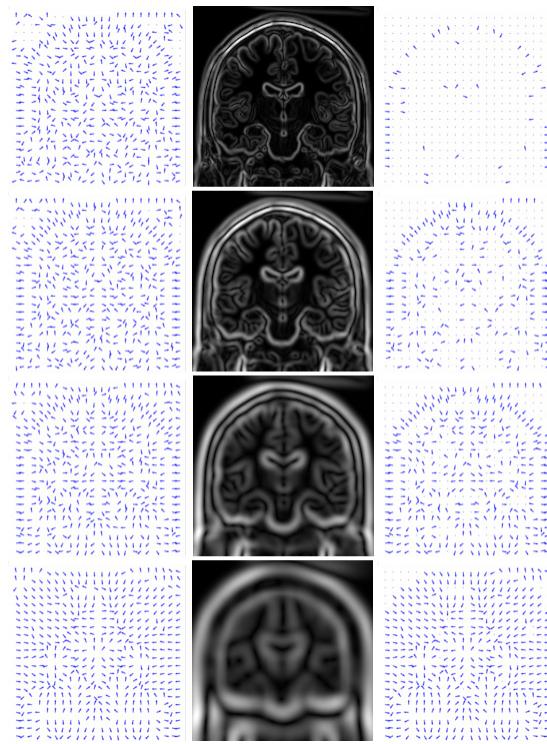


Fig. 3.3 Multi-scale pixel selection based on gradient magnitude threshold. Each row corresponds to a given differentiation kernel, where the top row corresponds to the finest scale and the bottom row to the coarsest. The left column shows all the computed gradient orientations, the middle column shows the gradient magnitude, and the right column shows the gradient orientations corresponding to selected locations. The percentage of selected pixels at each level is: 10.0%, 44.04%, 75.15% and 90.73%. Note that the multi-scale scheme does not involve a downsampling operation and simply modifies the scale (i.e. variance) of the Gaussian derivative kernel. The brain MRI image was obtained from BrainWeb [86].

scheme, then computing the image derivatives by convolution with linear operators, and then evaluating the gradient orientation. Consequently, the computational cost of a gradient orientation based metric which follows such strategy will be higher than that of a pixel-intensity based metric. Fortunately, one can adopt a computationally efficient approach for evaluating the gradient orientation of the deformed image with a linear mapping (known as the *pullback* or *covariant vector transformation law*) obtained with the derivative of the transformation function. Such a strategy has also been used in the context of DT-MRI registration [25], which is inherently in directional vector form.

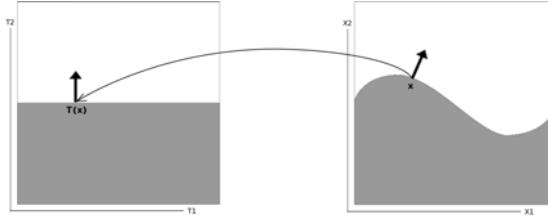


Fig. 3.4 Inversely Transformed Gradient Orientation. The left image corresponds to the moving image in its original coordinate space, the right image corresponds to the moving image evaluated on the inversely transformed coordinate space. Hence, a location, \mathbf{x} , on the transformed image coordinates (i.e. fixed image domain), corresponds to location $\mathbf{T}(\mathbf{x})$ on the original image coordinates (i.e. moving image domain).

The pullback differential is illustrated in Fig. 3.4, where the transformation function maps a location in the spatial domain of the fixed image, $\mathbf{x} = (x_1, \dots, x_D)$, to a location in the D -dimensional spatial domain of the moving image, $\mathbf{T} = (T_1, \dots, T_D)$. By expanding the derivative of the moving image, I_m , with respect to a particular dimension, x_j , of the fixed image spatial domain, one obtains the following expression

$$\frac{\partial(I_m \circ \mathbf{T})}{\partial x_j} \Big|_{\mathbf{x}} = \sum_i^D \frac{\partial I_m}{\partial T_i} \Big|_{\mathbf{T}(\mathbf{x})} \cdot \frac{\partial T_i}{\partial x_j} \Big|_{\mathbf{x}} \quad (3.11)$$

where the term $\frac{\partial T_i}{\partial x_j}$ corresponds to the (i, j) -th component of the spatial Jacobian matrix of the transformation function:

$$J_{\mathbf{T}} = \begin{bmatrix} \frac{\partial T_1}{\partial x_1} & \cdots & \frac{\partial T_1}{\partial x_D} \\ \vdots & \ddots & \vdots \\ \frac{\partial T_D}{\partial x_1} & \cdots & \frac{\partial T_D}{\partial x_D} \end{bmatrix}. \quad (3.12)$$

Re-arranging terms, the expression for the image gradient of the transformed moving image is,

$$\nabla I_m^\downarrow(\mathbf{x}) = J_{\mathbf{T}}^T(\mathbf{x}) \cdot \nabla I_m(\mathbf{T}(\mathbf{x})) \quad (3.13)$$

where $\nabla I_m(\mathbf{x}) = \left(\frac{\partial I_m(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial I_m(\mathbf{x})}{\partial x_D} \right)$ is the gradient of the original (undeformed) moving image. In other words, the gradient of the transformed moving image can be evaluated

as the product of the transposed spatial Jacobian matrix and the gradient of original (undeformed) image at the mapped location.

In the context of linear transformations,

$$\mathbf{T}(\mathbf{x}) = \mathbf{Ax} + \mathbf{b} \quad (3.14)$$

the expression in Equation 3.13 can be directly evaluated as,

$$\nabla I_m^\downarrow(\mathbf{x}) = \mathbf{A}^T \cdot \nabla I_m(\mathbf{T}(\mathbf{x})) \quad (3.15)$$

since \mathbf{A} is already known.

Hence, in registration contexts involving linear transformations, the derivatives of both images need to be computed only once at a pre-processing stage, and transformed vectors involve a left-side matrix multiplication with a known matrix. This allows the computational complexity of the method to be significantly reduced.

In order to demonstrate the computational advantage of evaluating the transformed gradient orientation with the pullback differential, a quantitative comparison of its computational cost is implemented in C++ with the Insight Toolkit library. Note that the conventional approach consists of the interpolation of the required voxel intensities followed by their convolution with a derivative operator, while the proposed approach consists of the interpolation of the original image gradient at the location of interest followed by a multiplication with the spatial Jacobian.

The benchmark involves the computation of the transformed image gradient with both approaches at 128,000 random locations for each of 100 random rigid transformations in a 2D and a 3D image. The average processing time is evaluated for each approach. Note that a linear interpolation scheme is employed for both intensity values and gradients, and a Gaussian derivative operator of varying radii ranging from 1 to 4 pixels is also employed. For the proposed approach, the initial computation of the original moving image gradient is not accounted for, since it is only done once prior to optimization.

The results are found in Table 3.1, which show the percentage of the average processing time of the proposed approach as compared to a conventional approach. Notice that one obtains a significant reduction in computational time for all radii. In particular, the smallest reduction is obtained with a 3D image and a derivative operator of radius 1, where the

proposed approach took approximately one third of the time a conventional approach takes. As the radius is increased, there is an even more significant reduction in processing times.

Derivative Operator Radius	2D Image	3D Image
1	30.69%	33.23%
2	19.00%	20.48%
3	12.92%	15.63%
4	9.81%	12.74%

Table 3.1 Comparison of processing time for evaluating transform image gradients. The values found in the table (expressed in percentages) correspond to the ratio of the average processing time of the proposed approach for computing the transformed image gradient to the average processing time of a conventional approach. The conventional approach first interpolates all required image intensities and then computes each derivative by convolution with a derivative operator. Percentages values below one hundred indicate faster processing times for the proposed approach.

3.1.3 Algorithmic Description

This section provides an algorithmic description of the proposed method, illustrated in Fig. 3.1, and also shown in pseudocode in Algorithm 1. The method is composed of two stages: a pre-processing stage and a multi-scale optimization stage. Furthermore, it requires the following configuration parameters: the matching criterion of the local measure, N ; the standard deviation of Gaussian derivative kernels at each scale; and a percentage, p , of high gradient magnitude locations at the finest scale of the fixed image, which effectively defines the threshold and hence sampling mask at such scale.

The pre-processing stage consists of the evaluation of image gradients of both images at all scales. The gradients are computed by convolution with corresponding discrete Gaussian derivative kernels. Notice that at different stages, the image size remains constant while the derivative kernel is changed. The sampling masks at each scale are automatically obtained by identifying their corresponding thresholds.

Once the pre-processing stage is concluded, a multi-scale optimization of the metric is performed. Starting with the coarser image scale, a subset of pixel locations whose gradient magnitude is larger than the given threshold is selected. The subset can be either the full sampling mask or a reduced pre-defined number of randomly selected locations within the

mask, where the latter option allows us to further reduce computational time but can potentially degrade registration robustness and accuracy.

The similarity metric is evaluated as the mean local measure of gradient orientation alignment on the selected subset and maximized with a covariance matrix adaptation evolution strategy [67], a non-gradient based optimizer. The result of the optimization is then used as the initial point for the next (finer) scale where the transformation is refined, and consequently for following scales. The final solution is obtained as the optimization result of the finest scale.

Algorithm 1 Proposed multi-scale registration algorithm.

```

1: procedure REGISTRATION( $I_f, I_m, p$ )
2:   for  $s \leftarrow 1, M$  do
3:      $\nabla I_f^s \leftarrow \text{gradient}(I_f, s)$ 
4:      $\nabla I_m^s \leftarrow \text{gradient}(I_m, s)$ 
5:   end for                                      $\triangleright$  Compute image gradients at each scale  $s$ .
6:    $\beta^M \leftarrow \text{percentile}(\nabla I_f^M, p)$        $\triangleright$  Find threshold at finest scale  $M$ 
7:   for  $s \leftarrow 1, M - 1$  do
8:      $\beta^s \leftarrow \frac{|\mathbf{K}^s|}{|\mathbf{K}^M|} \cdot \beta^M$ 
9:   end for                                      $\triangleright$  Find thresholds at coarser scales
10:   $\mathbf{T} \leftarrow \mathbf{I}$                                  $\triangleright$  Initialize transformation to identity function.
11:  for  $s \leftarrow 1, M$  do
12:     $\Omega \leftarrow \text{selectPixels}(\nabla I_f^s, \beta^s)$      $\triangleright$  Select a subset of pixels such that  $\nabla I_f^s > \beta^s$ 
13:     $\mathbf{T} \leftarrow \arg \max_{\mathbf{T}} S(\nabla I_f^s, \nabla I_m^s, \mathbf{T}, \Omega)$            $\triangleright$  Optimize metric
14:  end for
15:  return  $\mathbf{T}$ 
16: end procedure

```

The method has been implemented in both a CPU-based C++ application, and a hardware accelerated GPU-based application. The CPU-based application was developed as an extension of *Elastix* [85] which itself relies on the *Insight Toolkit* (ITK) Library [74]. Section 4.6.1 outlines the specifications of the GPU-based implementation, which had to be adjusted so as to fully exploit the massive parallel computational capacity of such device.

3.2 Diffeomorphic Registration Based on Gradient Orientation Alignment and Boundary Proximity

The registration framework presented in the previous section was limited to contexts with rigid transformations. It is important to note that rigid transformations involve a compact set of transformation parameters with a global scope (i.e. a change in any of the transformation parameters affects all transformed locations). In other words, the displacement encountered at a given location does not have any local degree of freedom, and is instead directly coupled with the displacement encountered at all other locations. Such phenomenon somewhat relaxes the task of measuring local image similarity, since it is only necessary to define a similarity measure which, when evaluated over a series of locations of interest, can provide an accurate measure of global image similarity to discriminate over a compact set of transformation parameters.

Unfortunately, in contexts involving non-linear transformations, the adopted transformation model typically involves a very large set of transformation parameters with local scope. In other words, displacements encountered at a given region can be fully independent of displacements encountered at an alternative (distant) region. Hence, where in the rigid registration context, one could rely on all evaluated locations to provide discriminability over the entire and compact set of transformation parameters; in the non-linear registration context, a small subset of evaluated locations provides discriminability exclusively over a small subset of transformation parameters. As such, any proposed local similarity measure is faced with the task of providing improved *local* discriminability.

This section proposes a diffeomorphic non-linear registration framework, referred to as *SymBA* (*Symmetric Boundary Alignment*), which directly addresses the challenge of improved *local* discriminability by extending and adapting the components found in the rigid registration technique proposed in Section 3.1.1. In particular, the proposed non-rigid registration framework is based on a diffeomorphic transformation function and focuses on matching detected edges from both images by maximizing gradient orientation alignment and minimizing the distance to the nearest corresponding boundary. It is characterized by the following major components: an inference of the location of boundaries of interest in both images, a local similarity metric based on gradient orientation alignment and distance to the nearest inferred boundary, and a velocity field optimized in a coarse to fine strategy with a gradient ascent technique. The following section provides a detailed description of

the proposed technique and its components.

3.2.1 Methodology

The rigid registration technique described in Section 3.1 involved the use of a gradient magnitude threshold, where all locations with gradient magnitude above such threshold were considered valid locations for evaluating the similarity metric. The gradient magnitude was used as a selection criterion since locations with high gradient magnitude were shown to involve gradient orientations with reduced uncertainty (i.e. reduced variability with regards to the true underlying orientation). Unfortunately, locations with a high gradient magnitude do not necessarily provide a pixel compact localization of the underlying boundary of interest, and exposed image boundaries are instead commonly characterized by a *band* (i.e. stripe) of high gradient magnitude pixels spanning a few pixels in the normal direction of the underlying boundary.

In order to minimize the ambiguity in the location of image boundaries that arises with the exclusive use of gradient magnitude, the proposed approach relies on an explicit inference of boundary locations performed with either a conventional off-the-shelf edge detector (e.g. Canny) or a context specific edge detection strategy. It is important to note that selecting an optimal edge detector for a given task is not a focus of this thesis. One of the main advantages of using a dedicated edge detector is that it provides a pixel compact edge location. For example, some edge detection methods rely on non-maximum suppression to identify boundaries with a pixel resolution, and some methods can even provide sub-pixel boundary location accuracy.

Once the boundary locations have been identified, a distance map is computed where the Euclidean distance to the nearest boundary is evaluated at every voxel. Fig. 3.5 illustrates the locations that are detected by an edge operator applied to a MRI T1-weighted brain volume, as well as the corresponding distance map obtained from the detected edge locations.

Once a set of boundary locations, Ω , and a corresponding distance map, D , are computed for each image, a localized similarity metric can evaluate gradient orientation alignment and Euclidean distance to the nearest boundary. For a location, \mathbf{x}_f , in the fixed

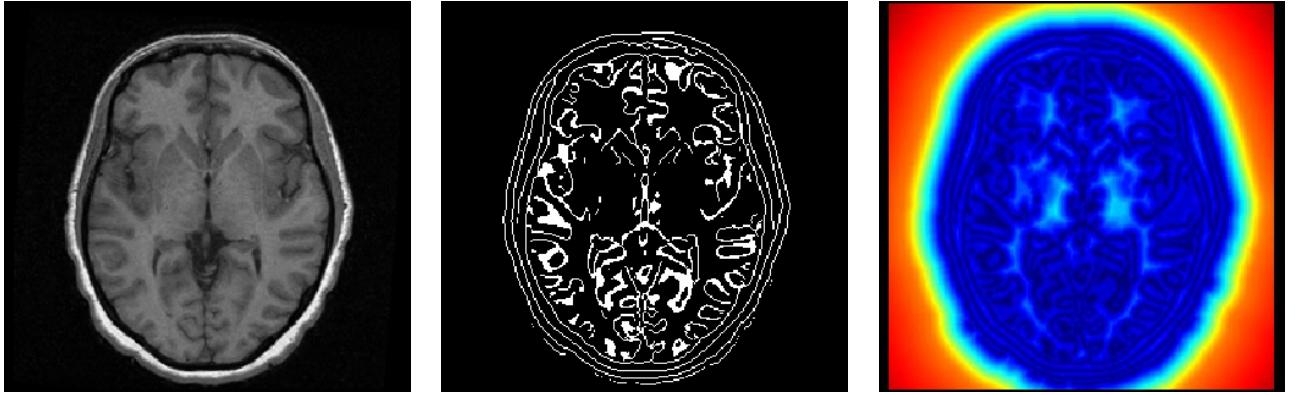


Fig. 3.5 Detected edges on a T1-weighted MRI brain volume. The first column shows the MRI brain volume. The second column shows the boundary locations detected with a Canny edge operator. The third column shows the distance image (jet colormap) computed from the detected edges.

image domain, the metric is expressed as,

$$s(\mathbf{x}_f; D_m^\downarrow, \nabla I_m^\downarrow) = \exp\left(-\frac{(D_m^\downarrow(\mathbf{x}_f))^2}{2\sigma^2}\right) \times \left\langle \frac{\nabla I_f(\mathbf{x}_f)}{|\nabla I_f(\mathbf{x}_f)|}, \frac{\nabla I_m^\downarrow(\mathbf{x}_f)}{|\nabla I_m^\downarrow(\mathbf{x}_f)|} \right\rangle^N \quad (3.16)$$

where ∇I_f is the fixed image gradient, $\nabla I_m^\downarrow = \nabla(I_m \circ \mathbf{T}_f)$ is the image gradient from the moving image deformed by the transformation function \mathbf{T}_f , and $D_m^\downarrow = D_m \circ \mathbf{T}_f$ is the distance map obtained from the moving image deformed by \mathbf{T}_f . Note that the transformation function, \mathbf{T}_f , is expressed with an f underscore so as to explicitly state that it involves a transformation from the fixed image domain to the moving image domain³.

The proposed similarity metric simply extends the measure of gradient orientation alignment, previously presented in Equation 2.6, with a measure based on the Euclidean distance to the nearest corresponding boundary. The main motivation behind such additional term is to provide the *local* discriminability required so as to resolve complex non-linear transformations. The similarity metric is characterized by two parameters: the standard deviation, σ , defining the falloff of the Gaussian function evaluating the distance to the closest boundary, and the gradient orientation selectivity, N , defining how gradient orientation alignment is evaluated and which follows the same rationale as the parameter found in Equation 2.6.

³ The transformation function, \mathbf{T} , used throughout the derivations in Section 3.1.1 also involved the mapping from the fixed image domain to the moving image domain. However, it was not necessary to explicitly express the distinction in such equations since it was the only transformation direction considered.

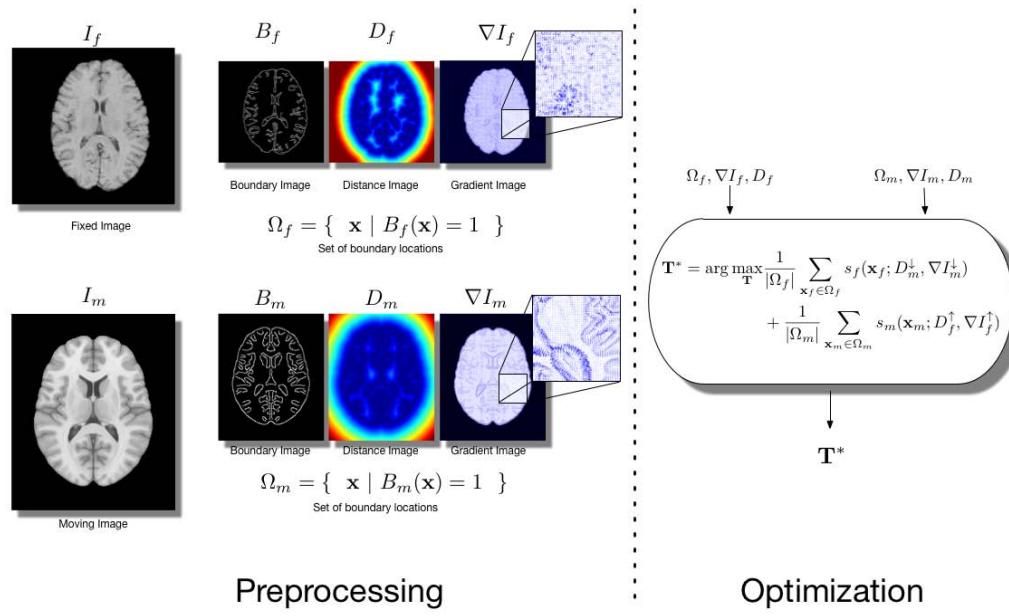


Fig. 3.6 Algorithmic outline of diffeomorphic registration technique, referred to as *SymBA*. The block diagram illustrates the pre-processing stage and optimization stage of the approach. The pre-processing stage involves the computation of image gradients, the inference of image boundary locations, and the computation of a distance map based on boundary locations. The optimization stage involves the maximization of gradient orientation alignment and minimization of distance to nearest boundary.

The rigid registration technique described in Section 3.1 involved an *asymmetric* voxel selection technique, where voxel locations identified in the fixed image domain are evaluated in terms of gradient orientation alignment at the spatially corresponding moving image location. Hence, such technique holds the assumption that locations selected from the fixed image domain correspond to an underlying anatomical boundary that is also exposed in the moving image domain, but that boundaries exposed in the moving image domain do not necessarily have a counterpart in the fixed image domain. This assumption is particularly relevant in multi-modal contexts where one image modality (e.g. MRI) exposes much more anatomical boundaries than a second modality (e.g. CT or US). However, in mono-modal contexts, both images expose the same anatomical boundaries and one can

safely assume that boundaries inferred in either image will indeed have a counterpart in the second image. As such, one can consider abandoning the asymmetric evaluation of the local metric. In particular, for mono-modal contexts one can define an energy function where the local similarity metric, defined in Equation 3.16, is evaluated over *both* sets of boundary locations,

$$S(\Psi_f, \Psi_m) = \frac{1}{|\Omega_f|} \sum_{\mathbf{x}_f \in \Omega_f} s(\mathbf{x}_f; D_m^\downarrow, \nabla I_m^\downarrow) + \frac{1}{|\Omega_m|} \sum_{\mathbf{x}_m \in \Omega_m} s(\mathbf{x}_m; D_f^\uparrow, \nabla I_f^\uparrow) \quad (3.17)$$

where $\Psi = (\nabla I, D, \Omega)$ is a notation convention used to group features of interest inferred from each image. Furthermore, $D_f^\uparrow = D_f \circ \mathbf{T}_m$ and $\nabla I_f^\uparrow = \nabla(I_f \circ \mathbf{T}_m)$ are the fixed distance map and gradient image deformed by $\mathbf{T}_m = \mathbf{T}_f^{-1}$. Note that the energy function involves a stronger notion of compromise between the inferences made in each image and that it requires the use of an invertible transformation.

The next section describes a proposed diffeomorphic transformation model which provides both a forward and backward mapping transformation, and which also allows for the computational efficient evaluation of related derivatives which are required to evaluate the proposed similarity metric and its gradient.

3.2.2 Diffeomorphism as the Integration of a Velocity Field

An energy function computed over both sets of locations involves the use of an invertible transformation. Specifically, the similarity metric for a boundary location, \mathbf{x}_m , in the moving image domain relies on the evaluation of $\mathbf{T}_m(\mathbf{x}_m)$. Hence, one requires a transformation that allows a forward mapping, $\mathbf{T}_f(\mathbf{x}_f)$, and a backward mapping, $\mathbf{T}_m(\mathbf{x}_m)$, such that $(\mathbf{T}_m \circ \mathbf{T}_f)(\mathbf{x}_f) = \mathbf{x}_f$.

Many commonly used non-linear transformation models used in registration contexts, such as B-Splines and Thin Plate Splines, are limited to the characterization of a forward mapping transformation and do not provide (nor guarantee the existence) of an inverse transformation (i.e. backward mapping transformation). Hence, such transformation models are simply not well-suited for the proposed registration technique. However, there are diffeomorphic registration methods, such as the one presented in [10], which directly address the task of characterizing both a forward mapping transformation and a backward mapping transformation. Unfortunately, such methods commonly involve the evaluation of the

displacement field and the corresponding inverse displacement field across the entire image domain at each optimization iteration. In particular, at each optimization iteration of the method presented in [10], the forward mapping displacement field is first updated based on the gradient of the energy function, and an iterative algorithm is then used to identify the corresponding backward mapping displacement field. Such operations involve a significant computational overhead and have limited use in the proposed registration method, which is focused on a *sparse* set of locations and thus does not require the evaluation of the transformation function across the entire image domain throughout the optimization.

In order to address such limitations, the proposed framework relies on a diffeomorphic transformation model characterized as the integration of a velocity field (discretized in time and space). Specifically, the forward mapping transformation is obtained as the composition of τ single-time-step transformations,

$$\mathbf{T}_f(\mathbf{x}_f) = \left(\prod_{t=1}^{\tau} \phi_{f,t} \right) (\mathbf{x}_f) = (\phi_{f,1} \circ \dots \circ \phi_{f,\tau}) (\mathbf{x}_f) \quad (3.18)$$

where the single-time-step transform at time index t is characterized as a displacement field, \mathbf{d}_t , evaluated based on the linearization of a velocity field, \mathbf{v} ,

$$\phi_{f,t}(\mathbf{x}_f) = \mathbf{x}_f + \mathbf{d}_t(\mathbf{x}_f) \quad (3.19)$$

$$= \mathbf{x}_f + \mathbf{v}_t(\mathbf{x}_f)\Delta t \quad (3.20)$$

where $\tau \times \Delta t = 1$.

The backward mapping transformation can be directly evaluated by integrating the velocity field in the negative direction,

$$\mathbf{T}_m(\mathbf{x}_m) = \left(\prod_{t=1}^{\tau} \phi_{m,t} \right) (\mathbf{x}_m) = (\phi_{m,1} \circ \dots \circ \phi_{m,\tau}) (\mathbf{x}_m) \quad (3.21)$$

where $\phi_{m,t}(\mathbf{x}_m) = \mathbf{x}_m - \mathbf{v}_t(\mathbf{x}_m)\Delta t$.

The main advantage of using such transformation model is that it can characterize both a forward and backward mapping transformation, which can be evaluated *on-demand* at specific locations of interest. Hence, each optimization iteration of the proposed registration method can focus exclusively at the locations where the similarity metric will be evaluated.

Interpolation

The transformation model adopted in the proposed framework employs a coarsely discretized velocity field image (with a voxel resolution coarser than the one of the images being registered). Hence, the evaluation of the velocity field across the image domain requires an interpolation function. Such dependency to an interpolation function is explicitly expressed in all following derivations. In particular, a forward mapping transformation involving τ time steps can be expressed as,

$$\mathbf{T}_f(\mathbf{x}_f) = \mathbf{x}_f + \sum_{t=1}^{\tau} \begin{bmatrix} \langle \mathbf{w}(\phi_{t-1}(\mathbf{x}_f)), \mathbf{V}_{t,1} \rangle \\ \langle \mathbf{w}(\phi_{t-1}(\mathbf{x}_f)), \mathbf{V}_{t,2} \rangle \\ \langle \mathbf{w}(\phi_{t-1}(\mathbf{x}_f)), \mathbf{V}_{t,3} \rangle \end{bmatrix} \Delta t \quad (3.22)$$

$$= \mathbf{x}_f + \sum_{t=1}^{\tau} \begin{bmatrix} \langle \mathbf{W}_t, \mathbf{V}_{t,1} \rangle \\ \langle \mathbf{W}_t, \mathbf{V}_{t,2} \rangle \\ \langle \mathbf{W}_t, \mathbf{V}_{t,3} \rangle \end{bmatrix} \Delta t \quad (3.23)$$

where \mathbf{w} represents an interpolation function, which maps a coordinate location to a vector of interpolation coefficients, \mathbf{W} . Furthermore, $\mathbf{V}_{t,d}$ is a vector composed of the velocity image pixel values for dimension, d , and at time index t .

In order to incorporate such transformation model in the proposed registration framework, it is necessary to derive the expressions which characterize the mapping of image gradient orientations from one domain to another, as well as the expressions which characterize the gradient of the transformation in terms of its parameters. Such derivations are covered in the following section.

Spatial Jacobian Matrix

This section outlines the derivation of the spatial Jacobian matrix of the proposed transformation function, which constitutes a key component for mapping image gradient orientations from one image domain to another, as well as a key component for the derivations of the gradient of the similarity function in terms of the transformation parameters.

The forward mapping transformation function is defined in Equation 3.18 as the com-

position of τ single-time-step transformations. Hence, one can rely on the chain rule of differentiation to express the Jacobian matrix of the transformation function as the product of the Jacobians of all τ single-time-step transformations. In particular, the Jacobian matrix of the forward mapping transformation can be expressed as,

$$J_{\mathbf{T}_f}(\mathbf{x}_f) = \prod_{t=1}^{\tau} J_{\phi_{f,t}}(\phi_{f,t-1}) \quad (3.24)$$

where $J_{\phi_{f,t}}(\phi_{f,t-1})$ is the Jacobian matrix of the single-time-step transformation encountered at time index t and at location $\phi_{f,t-1}$.

The expression for $J_{\phi_{f,t}}(\phi_{f,t-1})$ can be decomposed into two terms,

$$J_{\phi_{f,t}}(\phi_{f,t-1}) = \mathbf{Id} + J_{\mathbf{d}_t}(\phi_{f,t-1}) \quad (3.25)$$

where $J_{\mathbf{d}_t}(\phi_{f,t-1})$ is the Jacobian of the *displacement* \mathbf{d}_t encountered at time index t , which is expressed as,

$$J_{\mathbf{d}_t}(\phi_{f,t-1}) = \begin{bmatrix} \left\langle J_{\mathbf{W}}(\phi_{f,t-1}), \mathbf{V}_{t,1} \right\rangle \\ \left\langle J_{\mathbf{W}}(\phi_{f,t-1}), \mathbf{V}_{t,2} \right\rangle \\ \left\langle J_{\mathbf{W}}(\phi_{f,t-1}), \mathbf{V}_{t,3} \right\rangle \end{bmatrix} \Delta t \quad (3.26)$$

where $J_{\mathbf{W}}(\phi_{f,t-1})$ is the Jacobian of the vector of interpolation weights in terms of the location $\phi_{f,t-1}$ and can be easily evaluated in commonly used interpolation techniques, such as linear interpolation.

This result will be employed for subsequent derivations. In particular, the next section presents the required derivations for mapping vectors from one image domain to another in terms of the proposed transformation model, which constitute a fundamental component for the evaluation of gradient orientation alignment.

3.2.3 Transformation of Image Gradients

Section 3.1.2 illustrated that the pullback transformation of image gradients involves the left-side multiplication of a moving image gradient with the transpose of the Jacobian of the transformation at the point of interest. Such result led to a computationally efficient solution for the rigid registration method described in Section 3.1.3, where the Jacobian

matrix is constant at all points in the domain and is known. Unfortunately, the same strategy cannot be efficiently adopted in contexts involving non-linear transformations since the Jacobian matrix is no longer constant across the domain. This section describes a computationally efficient approach for estimating a transformed image gradient orientation which is specifically suited for the diffeomorphic transformation model employed.

The result from Section 3.1.2 is a particular example of a more general principle of the vector transformation law under a spatial transformation function. In particular, a vector, \mathbf{n}_y , defined on the transformed coordinate space (e.g. moving image domain), can be mapped back (“pulled-back”) so as to obtain the corresponding vector, \mathbf{n}_x , in the original coordinate space (e.g. fixed image domain),

$$\mathbf{n}_x = J_{\mathbf{T}_f}^T(\mathbf{x}) \cdot \mathbf{n}_y. \quad (3.27)$$

where $J_{\mathbf{T}_f}^T(\mathbf{x})$ is the transpose of the Jacobian matrix of the forward mapping transformation function, \mathbf{T}_f , at location, \mathbf{x} .

The Jacobian matrix is commonly expressed in terms of first-order partial derivatives,

$$J_{\mathbf{T}} = \begin{bmatrix} \frac{\partial T_1}{\partial x_1} \cdots \frac{\partial T_1}{\partial x_D} \\ \vdots \cdots \vdots \\ \frac{\partial T_D}{\partial x_1} \cdots \frac{\partial T_D}{\partial x_D} \end{bmatrix}. \quad (3.28)$$

However, it can also be expressed in terms of directional derivatives. For example, a Jacobian matrix for a 3D domain can be expressed as,

$$\begin{bmatrix} | & | & | \\ \frac{\partial \mathbf{T}}{\partial \mathbf{e}_1} & \frac{\partial \mathbf{T}}{\partial \mathbf{e}_2} & \frac{\partial \mathbf{T}}{\partial \mathbf{e}_3} \\ | & | & | \end{bmatrix} \quad (3.29)$$

where \mathbf{e}_i is the i -th basis of the fixed image coordinate space, such that $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$ and $\mathbf{e}_3 = (0, 0, 1)$.

Hence, the estimation of a Jacobian matrix can be decomposed into the estimation of a series of directional derivatives. For improved readability, the ongoing analysis is hereafter restricted to a 3D domain, which constitutes a main context of interest. In such context,

the Jacobian matrix is composed of three directional derivatives.

A general and straightforward approach which can be adopted for any transformation function is to estimate the directional derivatives of the Jacobian by a finite-difference method and consequently perform the matrix vector multiplication. The simplest finite-difference methods are the forward, backward and central-difference methods, which involve two function evaluations. One can further simplify such computation by fixing one of the two function evaluations at the point, \mathbf{x} , which was already required. Specifically, to minimize computational cost, each of these directional derivatives can be estimated with a forward-difference strategy, in the following fashion,

$$\frac{\partial \mathbf{T}}{\partial \mathbf{e}_i} \approx \frac{\mathbf{T}(\mathbf{x} + h \cdot \mathbf{e}_i) - \mathbf{T}(\mathbf{x})}{h} \quad (3.30)$$

where h is arbitrarily small and $\mathbf{T}(\mathbf{x})$ was already required. Hence, each directional derivative only involves a single additional transformation evaluation at $\mathbf{x} + h \cdot \mathbf{e}_i$.

In other words, the mapping of a vector in the moving image domain to the fixed image domain (i.e. $\mathbf{n}_y \rightarrow \mathbf{n}_x$) can be performed by first evaluating three directional derivatives and then multiplying the obtained Jacobian matrix transpose with the moving domain vector. This constitutes a general solution, valid for all transformation functions and vectors, which unfortunately leads to the computational overhead involved with the evaluation of three additional transformed locations.

The previously described procedure is a *general* approach for estimating the Jacobian matrix of any given transformation function, which is consequently used for evaluating a transformed image gradient. The procedure can be simplified by considering two points that are specific to the proposed registration technique. The first point is that the proposed registration technique is exclusively interested in the orientation of vectors, and can safely ignore any variations in the magnitude of a vector induced by a transformation. The second point is that the diffeomorphic transformation function adopted in this work is characterized as the composition of a series of single-time-step transformations that allow for a computationally simple evaluation of the Jacobian matrix, as described in Section 3.2.2. The following section describes how both points lead to a computationally efficient approach for evaluating a transformed normalized vector.

Pushforward of a Normalized Image Gradient

The result from Section 3.1.2 led to a *pullback* mapping of a vector in the transformed coordinate space to the original coordinate space (i.e. $\mathbf{n}_y \rightarrow \mathbf{n}_x$). This section describes an alternative *pushforward* mapping of a *normalized* vector in the fixed image domain to the moving image domain (i.e. $\mathbf{n}_x \rightarrow \mathbf{n}_y$ where $|\mathbf{n}_x| = 1$).

Consider the general pullback vector transformation expression,

$$\mathbf{n}_x = J_{\mathbf{T}_f}^T(\mathbf{x}) \cdot \mathbf{n}_y. \quad (3.31)$$

Restricting the vector \mathbf{n}_x to have a unit magnitude, one can define a rotation matrix, \mathbf{R} , which satisfies,

$$\mathbf{R} \cdot \mathbf{n}_x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \quad (3.32)$$

In particular, the rotation matrix, \mathbf{R} , is defined as,

$$\mathbf{R} = [\mathbf{n}_x \quad \mathbf{q}_x \quad \mathbf{r}_x]^T \quad (3.33)$$

where \mathbf{q}_x and \mathbf{r}_x are orthonormal vectors such that $\mathbf{n}_x = \mathbf{q}_x \times \mathbf{r}_x$.

Applying the rotation matrix to both sides of Equation 3.31 leads to,

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{R} \cdot J_{\mathbf{T}_f}^T(\mathbf{x}) \cdot \mathbf{n}_y \quad (3.34)$$

which can be re-arranged to

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} (J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{n}_x)^T \\ (J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{q}_x)^T \\ (J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{r}_x)^T \end{bmatrix} \cdot \mathbf{n}_y. \quad (3.35)$$

Note that the matrix $\Gamma = \begin{bmatrix} (J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{n}_x)^T \\ (J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{q}_x)^T \\ (J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{r}_x)^T \end{bmatrix}$ is composed of three directional derivatives

that are re-oriented so that the first derivative is evaluated in the direction of the vector \mathbf{n}_x . Furthermore, the equation can provide a direct expression for \mathbf{n}_y ,

$$\mathbf{n}_y = \boldsymbol{\Gamma}^{-1} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad (3.36)$$

which indicates that \mathbf{n}_y corresponds to the first column of $\boldsymbol{\Gamma}^{-1}$.

By re-arranging terms based on the components of $\boldsymbol{\Gamma}^{-1}$, the following analytical expression for \mathbf{n}_y is reached,

$$\mathbf{n}_y = \frac{1}{\det(\boldsymbol{\Gamma})} \left(J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{q}_x \times J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{r}_x \right). \quad (3.37)$$

Note that the proposed registration technique is exclusively interested in the *orientation* of the transformed vector and any scale terms (e.g. the magnitude of the determinant) can be safely ignored. The final expression of interest is thus,

$$\mathbf{n}_y \propto J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{q}_x \times J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{r}_x. \quad (3.38)$$

In other words, in a forward mapping transformation function, the fixed image gradient orientation can be expressed in the moving image domain as the cross product of two directional derivatives evaluated in directions orthogonal to \mathbf{n}_x .

The evaluation of the directional derivatives, $J_{\mathbf{T}}(\mathbf{x}) \cdot \mathbf{q}_x$ and $J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{r}_x$, can be performed with a minimal computational overhead in the proposed diffeomorphic transformation function. In particular, Section 3.2.2 described how the Jacobian matrix of the diffeomorphic transformation function is decomposed into the product of a series of Jacobians from single-time-step transformations. Hence, the left-side multiplication of a vector, \mathbf{q}_x , by the Jacobian matrix, can be re-expressed as,

$$J_{\mathbf{T}_f}(\mathbf{x}) \cdot \mathbf{q}_x = \left(\prod_{t=1}^{\tau} J_{\phi_t}(\phi_{t-1}) \right) \cdot \mathbf{q}_x \quad (3.39)$$

$$= J_{\phi_\tau}(\phi_{\tau-1}) \cdot \mathbf{q}_{x,\tau-1} \quad (3.40)$$

where $\mathbf{q}_{x,\tau-1}$ and “previous” vectors can be similarly defined as,

$$\mathbf{q}_{x,\tau-1} = J_{\phi_{\tau-1}}(\phi_{\tau-2}) \cdot \mathbf{q}_{x,\tau-2} \quad (3.41)$$

$$\mathbf{q}_{x,\tau-2} = J_{\phi_{\tau-2}}(\phi_{\tau-3}) \cdot \mathbf{q}_{x,\tau-3} \quad (3.42)$$

$$\vdots \quad (3.43)$$

$$\mathbf{q}_{x,1} = J_{\phi_1}(\phi_0) \cdot \mathbf{q}_{x,0} = J_{\phi_1}(\mathbf{x}) \cdot \mathbf{q}_x. \quad (3.44)$$

In other words, at every time index, t , an input vector, $\mathbf{q}_{x,t-1}$, is linearly mapped to a new vector, $\mathbf{q}_{x,t} = J_{\phi_t}(\phi_{t-1}) \cdot \mathbf{q}_{x,t-1}$. Thus, instead of explicitly computing the full Jacobian matrix, which involves τ matrix multiplications, and then performing a multiplication with \mathbf{q}_x , one can choose to ignore the explicit value of the full Jacobian matrix and directly perform τ linear mappings of \mathbf{q}_x at each time step. Note that the evaluation of the push forward transformed vector is performed in the same order of time integration as the evaluation of the transformed location and thus both operations can be coupled within the same computational procedure. Furthermore, such computations involve the evaluation of interpolation weights and their corresponding spatial derivatives which generally share many internal computations. In short, at every time index, t , one can combine the evaluation of a transformed location and two linearly mapped vectors into a single computational method. The combined approach for evaluating the transformation induced to both a coordinate location, \mathbf{x} , and a normalized vector, \mathbf{n} , is presented in pseudocode in Algorithm 3.2.3.

This section presented the computational approach adopted by the proposed registration framework for evaluating transformed image gradient orientations in an efficient fashion, which allows the evaluation of the proposed similarity metric. However, the optimization strategy also relies on the gradient of the similarity function with regards to transformation parameters, and its corresponding derivations are covered in the following section.

3.2.4 Optimization

The proposed framework employs a direct gradient ascent strategy so as to seek the maximization of the similarity function, and where the transformation parameters updated at every iteration correspond to the pixel values of the velocity field image. In particular, the

Algorithm 2 Forward Transform Point and Normalized Vector

```

1: procedure FOWARDTRANSFORMPOINTANDVECTOR( $\mathbf{x}, \mathbf{n}$ )
2:
3:    $(\mathbf{q}, \mathbf{r}) \leftarrow \text{OrthonormalVectors}(\mathbf{n})$ 
4:   ▷ Compute orthonormal vectors which satisfy  $\mathbf{n} = \mathbf{q} \times \mathbf{r}$ 
5:    $\mathbf{x}' \leftarrow \mathbf{x}$ 
▷ Initialize  $\mathbf{x}'$ 
6:
7:   for  $i \leftarrow 1, \tau$  do
▷ Loop over  $\tau$  time steps
8:      $(\mathbf{x}', \mathbf{q}', \mathbf{r}') \leftarrow \text{TransformPointAndVectors}(\mathbf{x}', \mathbf{q}', \mathbf{r}', \mathbf{V}_i)$ 
9:   end for
10:
11:   $\mathbf{n}' \leftarrow \mathbf{q}' \times \mathbf{r}'$ 
12:
13:  return  $(\mathbf{x}', \mathbf{n}')$ 
14: end procedure
15:
16: procedure TRANSFORMPOINTANDVECTORS( $\mathbf{x}, \mathbf{q}, \mathbf{r}, \mathbf{V}$ )
17:
18:    $(\mathbf{W}, \mathbf{J}_{\mathbf{x}}(\mathbf{W})) \leftarrow \text{ComputeInterpolationWeightsAndJacobian}(\mathbf{x})$ 
19:
20:    $\mathbf{x}' \leftarrow \mathbf{x} + \Delta t \cdot \begin{bmatrix} \langle \mathbf{W}, \mathbf{V}_1 \rangle \\ \langle \mathbf{W}, \mathbf{V}_2 \rangle \\ \langle \mathbf{W}, \mathbf{V}_3 \rangle \end{bmatrix}$ 
21:
22:
23:    $\mathbf{q}' \leftarrow \mathbf{q} + \Delta t \cdot \begin{bmatrix} \langle \mathbf{J}_{\mathbf{x}}(\mathbf{W}), \mathbf{V}_1 \cdot q_1 \rangle \\ \langle \mathbf{J}_{\mathbf{x}}(\mathbf{W}), \mathbf{V}_2 \cdot q_2 \rangle \\ \langle \mathbf{J}_{\mathbf{x}}(\mathbf{W}), \mathbf{V}_3 \cdot q_3 \rangle \end{bmatrix}$ 
24:
25:    $\mathbf{r}' \leftarrow \mathbf{r} + \Delta t \cdot \begin{bmatrix} \langle \mathbf{J}_{\mathbf{x}}(\mathbf{W}), \mathbf{V}_1 \cdot r_1 \rangle \\ \langle \mathbf{J}_{\mathbf{x}}(\mathbf{W}), \mathbf{V}_2 \cdot r_2 \rangle \\ \langle \mathbf{J}_{\mathbf{x}}(\mathbf{W}), \mathbf{V}_3 \cdot r_3 \rangle \end{bmatrix}$ 
26:
27:   return  $(\mathbf{x}', \mathbf{q}', \mathbf{r}')$ 
28: end procedure

```

k -th optimization iteration can be expressed as,

$$\mathbf{V}^k \leftarrow \mathbf{V}^{k-1} + \lambda^k \cdot \frac{dS}{d\mathbf{V}} \quad (3.45)$$

where $\frac{dS}{d\mathbf{V}}$ corresponds to the *total* derivative of the similarity function with respect to the velocity field pixel values, \mathbf{V} , and is evaluated as the average value of the total derivative of the local similarity metric, $\frac{ds}{d\mathbf{V}}$, at each evaluated location.

The total derivative $\frac{ds}{d\mathbf{V}}$ can be expressed as,

$$\frac{ds}{d\mathbf{V}} = \nabla_{\mathbf{T}_f}(s) \cdot \frac{d\mathbf{T}_f}{d\mathbf{V}} \quad (3.46)$$

where $\nabla_{\mathbf{T}_f}(s)$ is the gradient of the local similarity function, s , with respect to \mathbf{T}_f and $\frac{d\varphi}{d\mathbf{V}}$ is the total derivative of \mathbf{T}_f with respect to \mathbf{V} .

Note that the gradient $\nabla_{\mathbf{T}_f}(s)$ is composed of three partial derivatives (for a 3D domain),

$$\nabla_{\varphi}(s) = \left(\frac{\partial s}{\partial T_{f,1}}, \frac{\partial s}{\partial T_{f,2}}, \frac{\partial s}{\partial T_{f,3}} \right) \quad (3.47)$$

where $\frac{\partial s}{\partial T_{f,i}}$ corresponds to the partial derivative of the similarity function, s , with respect to the i -th component of the transformed coordinate location, \mathbf{T}_f . In this work, each partial derivative is estimated with a finite difference method.

The expression for $\frac{d\mathbf{T}_f}{d\mathbf{V}}$ can be derived by considering the dependency of the transformation function to the previous single-time-step transformation,

$$\mathbf{T}_f = \phi_{\tau} \quad (3.48)$$

$$= \phi_{\tau-1} + \mathbf{d}_{\tau} \quad (3.49)$$

$$= \phi_{\tau-1} + \begin{bmatrix} \langle \mathbf{W}_{\tau}, \mathbf{V}_1 \rangle \\ \langle \mathbf{W}_{\tau}, \mathbf{V}_2 \rangle \\ \langle \mathbf{W}_{\tau}, \mathbf{V}_3 \rangle \end{bmatrix} \Delta t \quad (3.50)$$

$$= \phi_{\tau-1} + \begin{bmatrix} \langle \mathbf{w} \circ \phi_{\tau-1}, \mathbf{V}_1 \rangle \\ \langle \mathbf{w} \circ \phi_{\tau-1}, \mathbf{V}_2 \rangle \\ \langle \mathbf{w} \circ \phi_{\tau-1}, \mathbf{V}_3 \rangle \end{bmatrix} \Delta t \quad (3.51)$$

where $\mathbf{T}_f = \phi_{\tau}$ since the transformation corresponds to the end time point of the integration of the velocity field, and \mathbf{V}_d is the vector of velocity field image pixel values corresponding to the spatial dimension d .

The differentiation of Equation 3.51 leads to the expression for the total derivative,

$$\frac{d\phi_\tau}{d\mathbf{V}} = \frac{d\phi_{\tau-1}}{d\mathbf{V}} + J_{\mathbf{d}_\tau}(\phi_{\tau-1}) \cdot \frac{d\phi_{\tau-1}}{d\mathbf{V}} + J_{\mathbf{d}_\tau}(\mathbf{V}) \quad (3.52)$$

$$= J_{\mathbf{d}_\tau}(\mathbf{V}) + \left(\mathbf{Id} + J_{\mathbf{d}_\tau}(\phi_{\tau-1}) \right) \frac{d\phi_{\tau-1}}{d\mathbf{V}} \quad (3.53)$$

where $J_{\mathbf{d}_\tau}(\phi_{\tau-1})$ is the Jacobian of \mathbf{d}_τ with respect to $\phi_{\tau-1}$, defined in Equation 3.26, $J_{\mathbf{d}_\tau}(\mathbf{V})$ is the Jacobian of \mathbf{d}_τ with respect to the velocity image pixel values, and \mathbf{Id} is the identity matrix.

The Jacobian $J_{\mathbf{d}_\tau}(\mathbf{V})$ essentially involves a sparse matrix whose non-zero components correspond to interpolation coefficients multiplied by Δt . For example, if the vector of velocity image pixel values is flattened into a linear vector (e.g. $\mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3)$) which concatenates the pixel values for all three components, one can express the matrix $J_{\mathbf{d}_\tau}(\mathbf{V})$ as,

$$J_{\mathbf{d}_\tau}(\mathbf{V}) = \begin{bmatrix} \mathbf{W}_\tau, \dots, 0 & 0, \dots, 0 & 0, \dots, 0 \\ 0, \dots, 0 & 0, \dots, \mathbf{W}_\tau, \dots, 0 & 0, \dots, 0 \\ 0, \dots, 0 & 0, \dots, 0 & 0, \dots, \mathbf{W}_\tau \end{bmatrix} \Delta t. \quad (3.54)$$

The expression found in Equation 3.53 can be simplified by adopting the assumption that $J_{\mathbf{d}_\tau}(\phi_{\tau-1}) \ll \mathbf{Id}$. The assumption corresponds to stating that local variations on displacements found in single-time-step transformations are relatively small, and play a negligible effect with relation to the identity matrix. It is important to note that $J_{\mathbf{d}_\tau}(\phi_{\tau-1})$ is proportional to Δt , and such assumption can be guaranteed with a sufficiently large number of time steps. Under such assumption the total derivative can be approximated as,

$$\frac{d\phi_\tau}{d\mathbf{V}} \approx J_{\mathbf{d}_\tau}(\mathbf{V}) + \frac{d\phi_{\tau-1}}{d\mathbf{V}} \quad (3.55)$$

which is a much simpler computation that simply accumulates the Jacobian across time indices.

The proposed registration framework adopts either a time-dependent/dynamic velocity field where the velocity field encountered at each time point is unique, or a time constant/static velocity field where the velocity field is common across all time points. In the context of a time-constant/static velocity field, the expression for updating the velocity

field at the k -th optimization iteration is defined as,

$$\mathbf{V}^k \leftarrow \mathbf{V}^{k-1} + \lambda^k \cdot \sum_i^{\tau} J_{\mathbf{d}_i}(\mathbf{V}) \quad (3.56)$$

which simply accumulates the Jacobians $J_{\mathbf{d}_{\tau}}(\mathbf{V})$ from all time points.

Alternatively, in the context of a time-dependent/dynamic velocity field, each of the τ velocity fields must be updated at each optimization iteration. In particular, the expression for updating the velocity field with time index, t , and at the k -th optimization iteration is defined as,

$$\mathbf{V}_t^k \leftarrow \mathbf{V}_t^{k-1} + \lambda^k \cdot J_{\mathbf{d}_t}(\mathbf{V}_t). \quad (3.57)$$

Multi-Resolution Optimization

As mentioned in the previous section, the proposed registration technique relies on a gradient ascent approach for optimizing the similarity metric. Furthermore, the learning rate of the gradient ascent is continuously adapted with a “bold driver” strategy, and the optimization is stopped when either a minimum rate of change or a maximum number of iterations is reached.

Many registration techniques adopt a multi-resolution image pyramid, described in Section 2.1.8, for simplifying the registration task. The proposed approach also adopts a related coarse to fine optimization strategy in which both the resolution of the images being registered and the velocity image is increased from stage to stage. It is important to highlight that the resolution of the velocity field (i.e. the voxel spacing of the discrete image defining the velocity field) employed will not necessarily correspond to the resolution of the images being registered. In other words, one can evaluate a deformation field with a finer voxel resolution than the one found in the velocity field.

3.3 Summary

In this chapter, I have detailed the general framework regarding a multi-modal rigid registration technique and a diffeomorphic registration technique. Both approaches are based on the common principle of assessing image resemblance at sparse prudently selected locations based on gradient orientation alignment. However, in order to truly exploit the sparse and

directional nature of the features used, I have developed a series of contributions that lead to the improvements in registration performance that are desired.

In the context of rigid registrations, I devoted a significant analytical effort to justifying the use of the gradient magnitude as an indicator of the confidence on a given gradient orientation. The analysis was framed as a Bayesian inference problem and dependent on the linear operator used for computing the derivative. As such, it yielded the key conclusion that the confidence of gradient orientations can be compared across scales and a common confidence threshold can be defined for all scales. I also illustrated the importance of using the pullback differential for evaluating the transformed image gradient. These two points are critical for achieving a highly efficient performance, and the experimental results illustrating such performance are presented in Chapters 4 and 5.

In the context of non-rigid registrations, I presented an approach that explicitly relies on inferred boundaries and enforces multiple constraints on a diffeomorphic transformation function to yield a highly efficient registration performance. In particular, the approach defines a symmetric energy function in which boundaries inferred in each image are warped to maximize the gradient orientation alignment and minimize the distance to the nearest boundary at their mapped location. The framework characterizes a diffeomorphic transformation as the integration of a time-constant or time-dependent velocity field. Such setup allows the optimization to focus on locations where the similarity metric is being evaluated while preserving a parametric description of the transformation throughout the full image domain.

Chapters 6 and 7 demonstrate a series of experimental results highlighting the advantages brought forward by the proposed non-rigid registration technique. In particular, Chapter 6 illustrates how the proposed technique leads to significantly reduced processing times in the context of inter-patient non-linear registration of brain MRI volumes as evaluated with a publicly available dataset and compared against fourteen state-of-the-art registration techniques. Chapter 7 then evaluates the registration performance of the proposed technique in the context of patient to atlas brain registration for the generation of healthy tissue priors which consequently serve as inputs to automatic brain lesion detection and segmentation algorithms. In such context, the final lesion and detection performance is evaluated with a dataset of real clinical cases and it is shown that the tissue priors generated with the proposed registration technique can lead to slightly improved detection performance with highly reduced registration computation times. A critical advantage of

such reduction in processing times lies in large scale clinical trials that rely on medical image analysis pipeline where non-rigid registration technique typically involve a major component of the overall computational cost.

Chapter 4

Image Guided Neurosurgery: Clinical Context and Results

The previous chapter provided the algorithmic details regarding the proposed framework for image registration, with particular focus on yielding improved computational efficiency in medical imaging contexts with challenging image models. This chapter focuses on describing the critical relevance of the proposed registration technique in the context of neuro-surgical procedures that rely on image guidance, referred to as Image Guided Neurosurgery (IGNS) or Neuronavigation, and adopt intra-operative ultrasound (iUS) for improved intra-operative guidance. One of the major challenges for truly taking advantage of iUS lies in the fast and robust registration of pre-operative volumes (such as MRI) to the iUS volume so as to update all pre-operative data. This chapter demonstrates how the proposed registration technique (along with other technical contributions) can lead to a complete neuronavigation solution based on iUS with minimal interference in the operating room, which constitutes a proven framework that is unmatched by alternative commercial solutions.

The chapter is structured as follows. Section 4.1 briefly describes the clinical background regarding intra-cranial tumour resection and highlights the importance of improved intra-operative guidance so as to maximize the removal of pathological tissue and minimize the damage to eloquent tissue. Section 4.2 and 4.3 describes the technical challenges involved in neuronavigation and illustrates the particular appeal for adopting tracked iUS as a framework for intra-operative imaging. Section 4.3.2 provides a description of the unique difficulties that characterize this registration problem and also outlines previous

attempts that have addressed such context with limited success. Section 4.4 then presents a series of experimental results obtained from the proposed registration framework with the use of a publicly available dataset of *real* clinical cases. The demonstrated registration performance, with some additional technical contributions, ultimately illustrates how the proposed techniques can provide a significant clinical impact in this context.

4.1 Intra-cranial Tumour Resection

This section describes the clinical context of intra-cranial tumour resection, which constitutes a key neurosurgical procedure that can substantially benefit from improved intra-operative guidance so as to maximize clinical outcomes. Intra-cranial tumours include primary and secondary tumours in the brain with challenging clinical outcomes and where surgical resection is a common treatment. A key prognostic factor in neuro-oncology is believed to be the extent of surgical resection. Some studies [143] estimate the benefits of a total resection as extending the survival from around 11 to 14 months in glioblastoma and from around 60 to 90 months in low grade glioma. One of the main challenges in achieving a total resection lies in accurately identifying tumour tissue throughout the procedure, particularly with relation to eloquent tissue (i.e. healthy brain tissue related to critical functions). The primary objective of the surgical procedure is thus to maximize the removal of tumour tissue while minimizing the damage on healthy tissue.

Various intra-operative technological strategies have been proposed to improve the identification of residual tumour throughout the procedure. Such strategies typically involve a neuronavigation system, intra-operative imaging modalities, and possibly intra-operative contrast agents. Among the proposed intra-operative strategies, iUS has been heavily advocated [6, 17, 63, 77, 91, 92, 124, 142, 152, 165] as a minimal footprint, cost-effective and clinically valuable modality for neurosurgical procedures. As such, iUS can be considered as a direct alternative to intra-operative MRI or intra-operative CT, which may provide higher quality imaging but at the expense of substantial and costly modifications to the operating room. In order to illustrate the importance of incorporating an intra-operative imaging solution, the next section will first provide a brief technical description of the components involved in a conventional neuronavigation system and what kind of limitations they encounter if they are based exclusively on pre-operative imaging data.

4.2 Neuronavigation

Neuronavigation, or Image Guided Neurosurgery, is a commonly used framework with commercial solutions offered by companies such as MedtronicTM, BrainLabTM and StrykerTM. The main feature of a neuronavigation system, as shown in Fig. 4.1, is the integration of medical imaging data with a real-time 3D tracking technology, which allows the clinician to visualize the location of a 3D tracking tool overlayed onto the medical image. The technological framework can be considered the surgical analogy of a GPS system which continuously provides an indication of someone's location and overlays it in a street map. Common 3D tracking technologies found in neuronavigation systems are either based on stereo infrared optics or electromagnetic tracking. Optic based tracking technology is significantly more common and involves the use of a stereo infrared camera in conjunction with reflective infrared spheres which serve as markers.

A brain tumour resection procedure with support from a neuronavigation system involves a series of stages that are depicted in Fig. 4.2. The first stage involves the acquisition of pre-operative images, the diagnosis of brain tumour and the planning of the surgical procedure. For example, the clinical team can choose to acquire a series of MRI volumes so as to identify the location and size of the brain tumours. The surgical team may then choose to manually segment the tumour tissue and other anatomical regions of interest so that they can be displayed as unique 3D objects in the neuronavigation system. The clinicians are then faced with the task of defining the surgical plan to be followed in the procedure which involves: defining location and size of the craniotomy, defining the trajectory or path to reach the tumour, and identifying the eloquent regions that are in danger of being damaged by the procedure.

The second stage relates to the activities that take place in the operating room during the procedure. In order to setup the neuronavigation system in the operating room, the pre-operative image data must be first loaded into the system. Then, the clinician must register the coordinate space of the real-world, referred to as "world space", to the coordinate space of the image data, referred to as "image space". A common approach to resolve such registration involves the manual identification of six to eight homologous landmarks in the patient's head (e.g. tip of the nose, ears, etc.) with a tracked pointing tool and correspondingly in the pre-operative image. Fig. 4.3 illustrates a common setup used for performing a landmark based registration. Once the real world is registered to the pre-

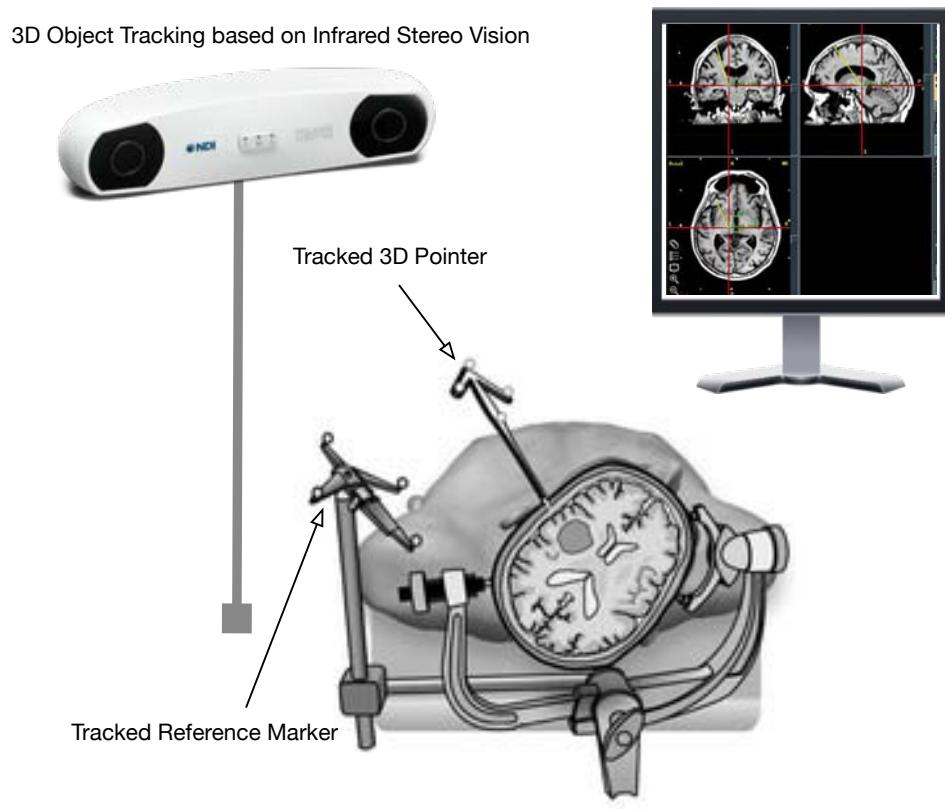


Fig. 4.1 Neuronavigation with tracked 3D pointer. A pointing tool with reflective infrared spheres can be continuously tracked in 3D with the use of a stereo infrared camera. A reference marker (also with reflective spheres) may also be used so as to define an arbitrary coordinate reference frame in relation to the reference marker position. Image obtained from [92]

operative image(s), the clinician can use the 3D tracked pointer to probe different points in the patient's brain and visualize its corresponding location in the pre-operative image, such as illustrated in Fig. 4.1.

The final stage involves a post-operative analysis of the extent of tumour resection. In particular, a series of post-operative images are acquired and are consequently analyzed by the clinical team so as to evaluate the extent of the tumour removed and if there was significant damage to nearby healthy regions.

The neuronavigation system described so far is restricted to mapping a real world 3D location to pre-operative image data. As such, it only provides a spatially registered pre-

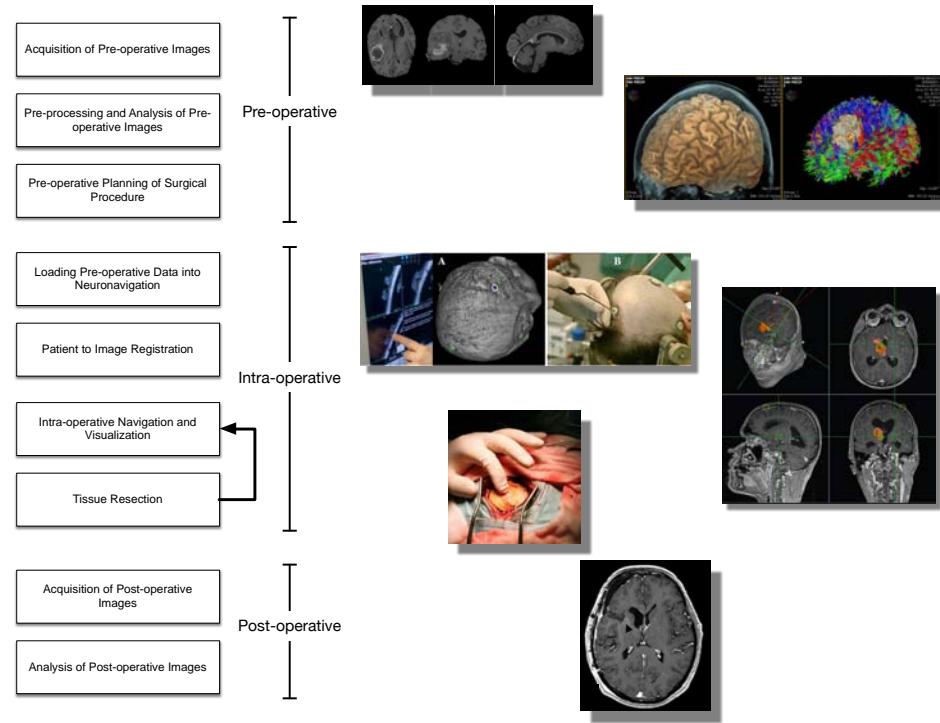


Fig. 4.2 Flowchart of a surgical procedure based on neuronavigation. Images obtained from [92], <http://www.synaptivemedical.com>, and <http://martinos.org/qtim/miccai2013>.

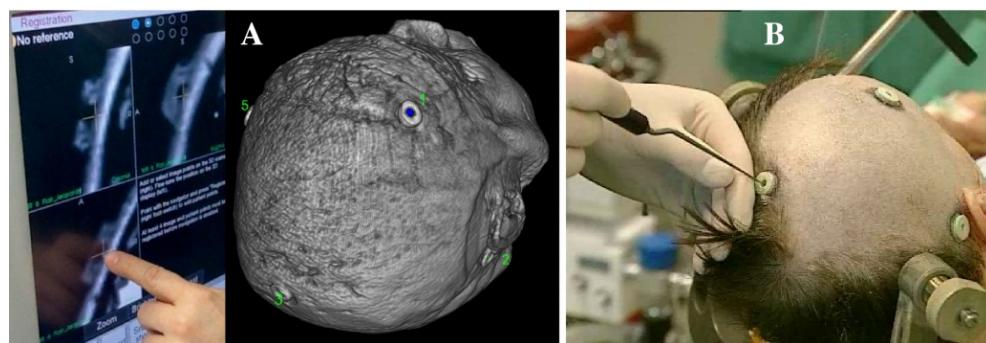


Fig. 4.3 Patient registration to image space based on homologous landmarks. Image obtained from [92].

resection snapshot of the patient's anatomy, and any anatomical changes that occur any time after the acquisition of the pre-operative images and particularly throughout the proce-

dure can simply not be captured. Unfortunately, brain soft-tissue deformations, commonly referred to as brain-shift, are a common occurrence in neurosurgical procedures involving brain tumours and automatically lead to a degradation in navigation accuracy. They are brought forward by the treatment of intra-cranial pressure prior and during surgery and by external forces applied during surgery, such as that related to retractors and gravity. Furthermore, deformations brought forward by tumour tissue resection and draining associated with cysts additionally lead to even more significant anatomical displacements. Fig. 4.4 illustrates the prominent swelling of brain tissue brought forward by changes in intra-cranial pressure in a brain tumour resection surgery and Fig. 4.5 illustrates the phenomenon of brain-shift with corresponding pre and intra-operative MRI image slices.

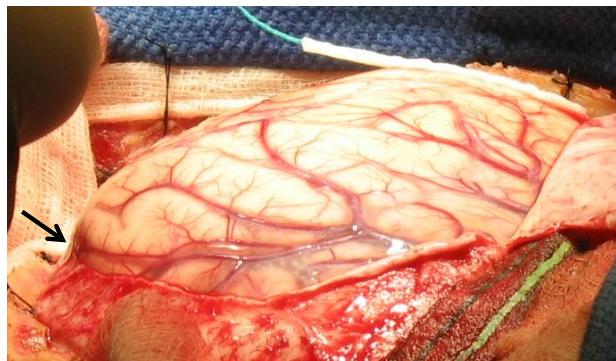


Fig. 4.4 Brain-shift caused by changes in intra-cranial pressure. Notice that the brain is bulging out of the dural opening, illustrated with a black arrow. Image obtained from [102].

The clear need for visualizing anatomical changes that occur throughout a brain tumour resection procedure leads to the adoption of an intra-operative modality. Intra-operative MRI and CT are commonly used modalities that can generate high quality images but at significant financial and practical expense. Alternatively, intra-operative US is a financially accessible and minimal footprint modality that can provide a more practical solution for intra-operative imaging. The following section describes the use of tracked iUS in neuron-navigation systems and the technical support it relies on.

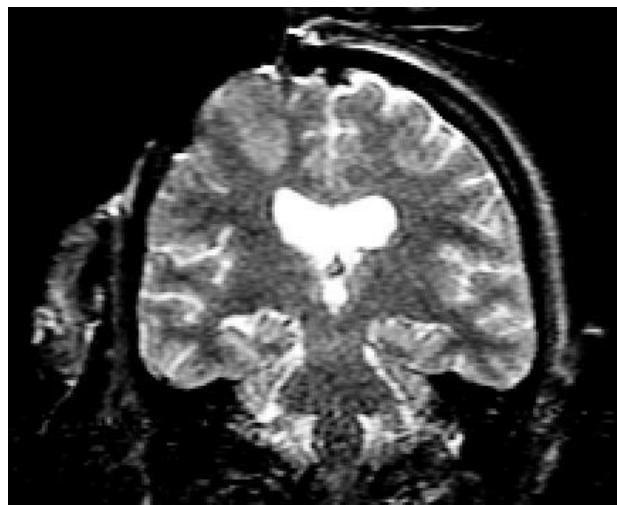


Fig. 4.5 Illustration of brain-shift encountered after removal of the dura and imaged with a T2-weighted intra-operative MRI. Notice how the deformation of the left ventricle follows the direction of the tumour bulging. Image obtained from [111].

4.3 Navigation with Tracked Ultrasound

Tracked iUS is a highly appealing imaging modality for clinical interventions. It is portable, free of health risks, relatively inexpensive, and allows for a simple integration in the operating room. Furthermore, it allows for real-time imaging and can provide continuous visual feedback throughout the intervention.

Fig. 4.6 illustrates the use of a tracked US in a neuronavigation system. The US probe is mounted on a tracked piece so that the probe's location and orientation can be identified by the tracking system. Modern US systems natively generate 2D images or 3D volumes. In the context of neurosurgery, 2D US systems are typically preferred since they involve probes of smaller dimensions that can be more easily positioned within a craniotomy.

One should note that it is necessary to perform an ultrasound image calibration so as to resolve the fixed relationship between the 3D location of the US probe and the location of its corresponding 2D image plane, as shown in Fig. 4.7. The problem of robustly and accurately resolving the spatial calibration of an ultrasound probe continues to be an important research topic [31, 73, 93, 106, 121], and is not covered in detail in this work.

Once the probe is properly calibrated, the neuronavigation system can track the location

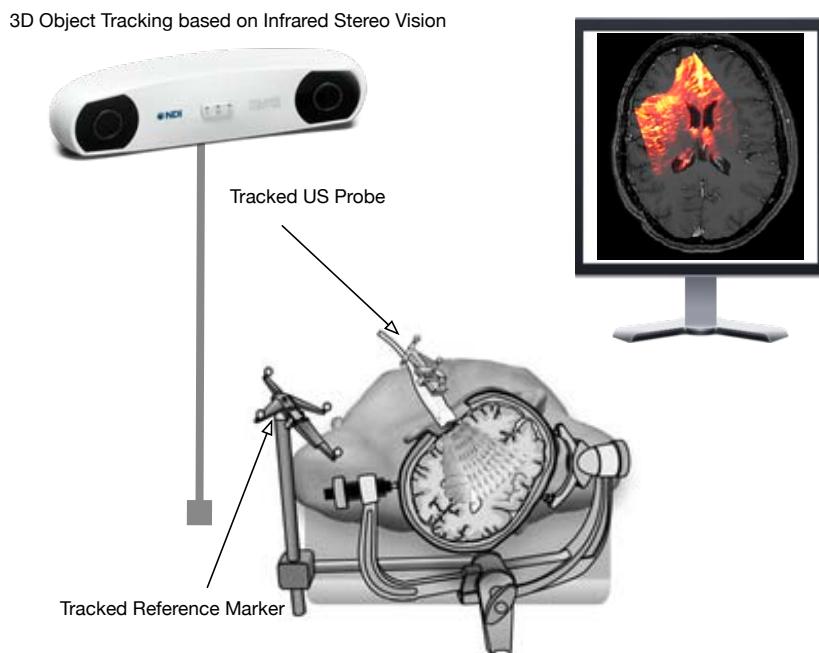


Fig. 4.6 Neuronavigation with tracked ultrasound. The setup involves a infrared stereo camera for continuously tracking the 3D location and orientation of a ultrasound probe with the use of reflective infrared spheres fixed to the probe. The acquired intra-operative US volume, illustrated with a heat map in the display, can then be accurately overlayed on the pre-operative MRI volumes. Image obtained from [92].

and orientation of each 2D image slice as it is being captured. The neuronavigation system can rely on a volume reconstruction algorithm [41, 61, 135, 136, 151] to generate a 3D volume from the sweep of tracked 2D image slices. Finally, the reconstructed 3D volume can be displayed in the neuronavigation system and visually compared with the pre-operative MRI volume.

In summary, the use of a tracked iUS involves the minimal requirements of properly mounting a known configuration of reflective spheres to the US probe and inferring the mapping between the position of the reflective spheres with relation to the location and orientation of the US image plane. However, the US device itself and the intra-operative procedure required to make use of it have practically negligible interference in the operating room in comparison with other modalities. Fig. 4.8 shows a complete commercial

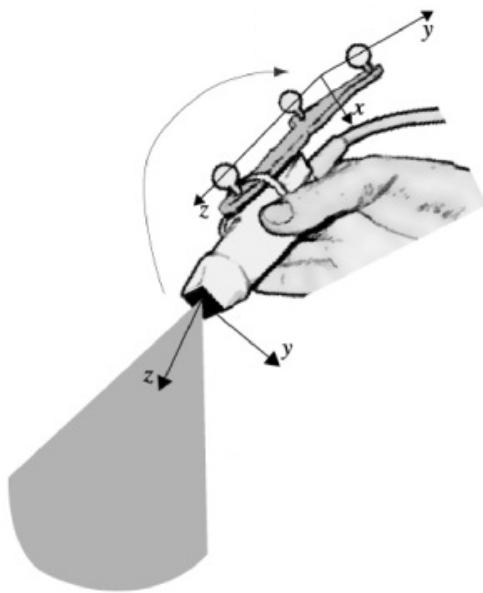


Fig. 4.7 Tracked ultrasound probe. Image obtained from [92].

neuronavigation system that includes support for tracked iUS. In direct contrast, Fig. 4.9 shows the proposed operating room configuration to allow integration of an intra-operative MRI. Note how the neuronavigation system with iUS support is highly portable and can be easily displaced throughout the procedure. Alternatively, the intra-operative MRI solution involves a custom built operating room configuration.

It is important to note that the practical advantages for the integration of iUS in the clinical procedure come with a significant penalty in imaging quality and consistency with relation to modalities like MRI or CT. Intra-operative US suffers from the challenge that the surgeon must be trained in the positioning and coupling of the US probe, as well as in the anatomical interpretation of the acquired images. Such US-specific training is typically limited to US radiologists, and not neurosurgeons. As such, there is generally an important learning curve associated with integrating iUS in the procedure. Furthermore, the technical characteristics of iUS lead to a restricted field-of-view with a trade-off between image resolution and penetration depth. Additionally, there are various US specific imaging artefacts (e.g. shadows, speckle), which can lead to a misinterpretation of the image if they

are not well understood. In direct contrast, competing modalities such as intra-operative MRI provide a far less restricted field of view and can be directly interpreted by a clinician (as any other MRI image). Even with all its imaging challenges, the highly reduced cost of iUS and its limited disturbance of the clinical procedure continue to make iUS a highly appealing image modality for many clinical interventions.



Fig. 4.8 Commercial neuronavigation system with support for iUS. Image obtained from <http://www.sonowand.com/>.

4.3.1 Navigation Accuracy

The clinical benefit of incorporating a neuronavigation system in a surgical procedure is directly related to the navigation accuracy that it can provide. A particularly important measure of navigation accuracy is the difference between the true anatomical location of a surgical tool and the anatomical location it is mapped to in the medical images. Unfortunately, it is not easy to characterize such accuracy in a clinical setting. Hence, most efforts [91, 125] have been limited to a controlled lab environment and the efforts [157]



Fig. 4.9 Operating room configuration integrating an intra-operative MRI. An example operating room based on Medtronic™Polestar N30. Image obtained from <http://www.medtronic.com/>.

that do characterize the accuracy in a clinical setting have been limited to rigid body assumptions.

In the context of neuronavigation with iUS, there is an important distinction regarding the error in accuracy relating to the mapping between pre-operative volumes and the patient space, and the error in accuracy relating to the mapping between intra-operative images and the patient space. The first source of error is associated to the precision of the patient to volume registration performed at the beginning of the procedure, which is typically affected by patient displacement throughout the procedure. In particular, some studies have shown [157] that there is a consistent decrease in such accuracy with relation to the duration of the procedure. Additionally, navigation accuracy can also be significantly affected by soft-tissue brain deformations, referred to as brain-shift [88, 125, 129]. Such soft-tissue deformations generally occur after the dura has been removed and are due to variations in intra-cranial pressure or external forces, such as gravity or retractors.

The second source of error is directly related to the precision of the iUS spatial calibration and the precision of the 3D tracking framework. The distinction between the two sources of navigation accuracy error is particularly important since it implies that intra-operative displacements can only affect the first type of error while having no effect on the second type of error. Hence, one can safely rely exclusively on a tracked iUS for navigation

and guidance with no concern regarding intra-operative displacements.

Unfortunately, there is typically a variety of pre-operative imaging data that bears critical relevance throughout the procedure and there is significant interest in ensuring that it is properly mapped to the patient's space. For example, consider a set of manual anatomical annotations performed on a pre-operative MRI which identify critical regions of eloquent tissue corresponding to high priority functions (e.g. speech). In the presence of intra-operative deformations, there is a clear interest in updating the location of such critical regions so as to fully ensure that they do not suffer any damage. In order to address this concern, the use of iUS can play key role where it can be used so as to *correct* for intra-operative displacements by registering the pre-operative image data to the iUS. This is a fundamental motivation behind the development of a multi-modal registration technique than can robustly resolve the spatial mapping between a pre-operative MRI brain volume and an iUS image. A successful and fast registration effectively allows the clinician to continue relying on all pre-operative imaging data once it has been updated to reflect the intra-operative deformations, and with no particular concern regarding a significant degradation in navigation accuracy. The next section briefly describes some of the previous work related to this task and consequently present a series of results with our proposed technique yielding sub-second processing times with robust performance.

4.3.2 Registration of Brain MRI to US

The registration of pre-operative MRI to iUS is a problem faced with multiple challenges brought forward by the widely different image formation models of each modality. MRI can be largely characterized as a tissue-type based modality, where the image intensity of a given voxel is mainly a function of the tissue found within the voxel's volume. On the other hand, US images illustrate different acoustic impedance transitions encountered by the ultrasonic wave. Figs. 4.10(a)-4.10(d) show four different neurosurgical cases and illustrate the differences between the two modalities. Notice that the MRI allows for an accurate identification of multiple soft-tissue types such as grey matter, white matter and bone, and hence also permits the localization of many anatomical structures on the brain (e.g. lateral ventricles, falx, sulci, grey and white matter boundaries, etc.). Alternatively, the corresponding US only exposes the tumour tissue and its boundary (with some degree of uncertainty), and also depicts part of some key structures like the falx and the lateral

ventricles. In short, while the MRI can expose a variety of anatomical structures in the brain, the iUS can only expose a few key anatomical boundaries and involves complex artefacts such as speckle and a non-homogeneous intensity response. Such phenomena are extremely challenging for any multi-modal image registration technique. In particular, the limited anatomical exposure depicted by iUS can be considered an example of complex image occlusion, where an image simply does not depict some components of the scene. Complex image occlusion in conjunction with a strongly non-homogeneous image formation model are a direct challenge for intensity-based similarity metrics, such as MI, that rely on the predictive power of a fixed global mapping of inter-modality intensity values.

In the particular context of brain tumours, we are also affected by the fact that pathologies from different cases can have quite unique image features in each modality. Figs. 4.10(a)-4.10(c) contrast three cases with significantly different depictions in each modality. In particular, we can observe that the inner tumour tissue and the tumour boundaries in the iUS are exposed with quite distinctive image characteristics in each case. For example, Fig. 4.10(c) depicts tumour tissue with very high US intensity values, but does not allow for an accurate identification of its boundary. Alternatively, in Fig. 4.10(b), the tumour tissue is depicted with a low US intensity and the boundary can be identified with increased certainty. Furthermore, Figures 4.10(b) and 4.10(c) provide a prominent depiction of the lateral ventricles in the iUS, a highly informative anatomical structure for identifying a match across modalities. On the other hand, Fig. 4.10(a) provides a very weak depiction of the lateral ventricles. This kind of variability in exposing anatomical structures in iUS is a clear challenge for conventional multi-modal registration approaches that assume a global hard mapping between image features of one modality to image features of a second modality.

The registration of MRI to US images has been previously addressed by various works [4, 40, 78, 87, 101, 105, 127, 128]. Some approaches [22, 131] rely on gradient magnitude as an image feature of interest in conjunction with a conventional multi-modal similarity metric such as mutual information (MI) or normalized cross-correlation (NCC). Other proposed techniques rely on local-phase [101] as a feature in conjunction with MI. A major challenge encountered by such approaches is that the image intensity response found in US is significantly non-homogeneous. Consider the coronal view found in Fig. 4.10(d) where we can clearly observe how the US intensity decays with relation to the distance from the probe. In particular, notice that the US pixel intensities corresponding to white

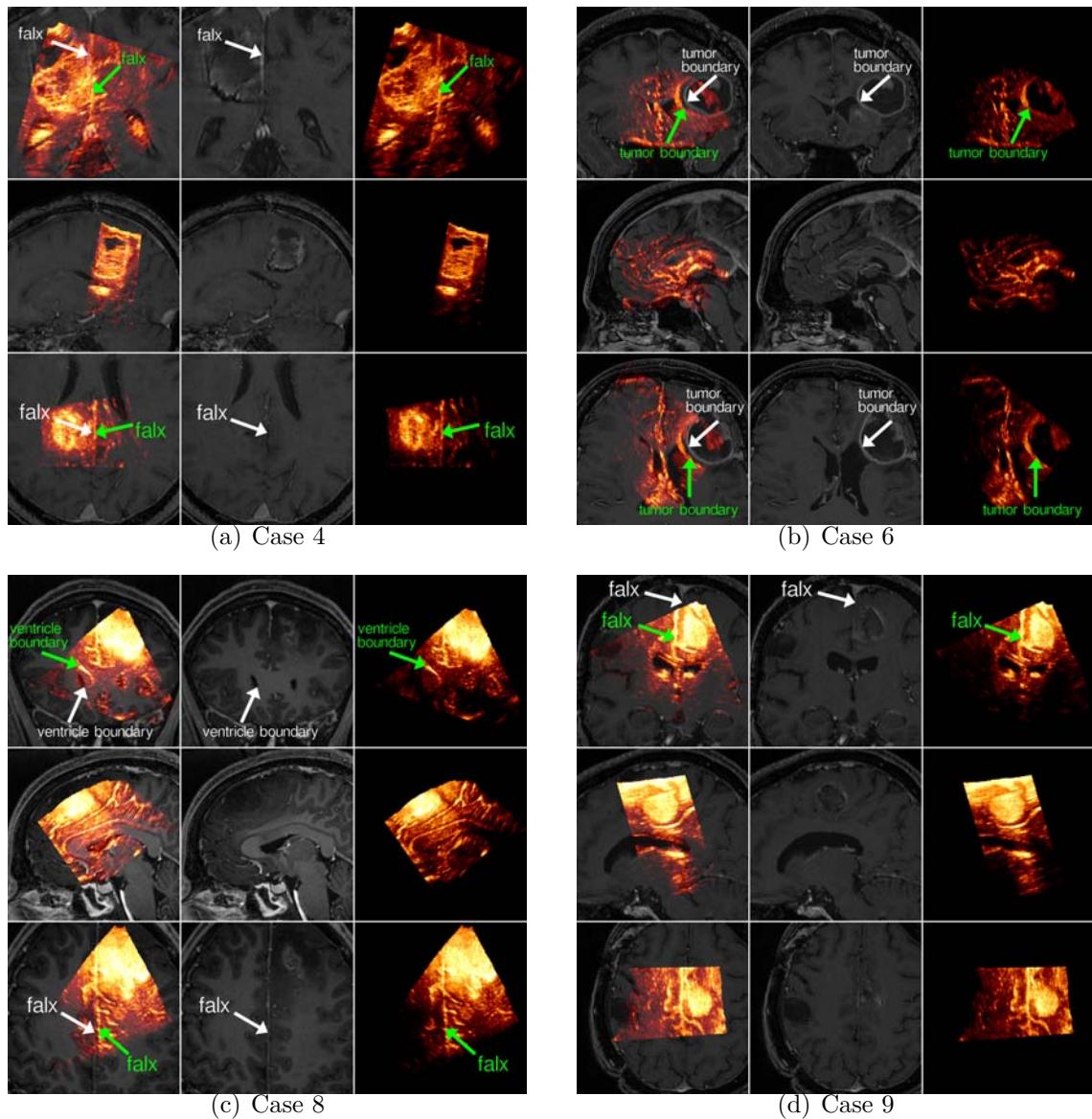


Fig. 4.10 Pre-operative MRI volume and the iUS volume in a brain tumour resection surgery as mapped by an initial patient to volume landmark registration. The first column shows the MRI image in greyscale with the corresponding iUS overlaid and heat map colored. The second column shows the MRI image. The third column shows the iUS. The top row shows a coronal view, the middle row shows a sagittal view, and the bottom row shows a transverse view. Anatomical structures found in the iUS are identified by a green arrow and label, while structures found in the MR are identified by a white arrow and label.

matter tissue are far from consistent, and will likely result in a degradation of registration performance when using a similarity metric that involves the full image domain. There have also been approaches [78, 87] that propose a preprocessing stage in which imaging artefacts (e.g. speckle, noise) are reduced and consequently register the preprocessed images with a multi-modal metric like normalized mutual information (NMI) evaluated over a sampling mask that typically covers the tumour volume and part of its surrounding region. Alternatively, other works [4, 105] have proposed the generation of a pseudo-US by hard-mapping segmented structures in the MRI to specific intensity values, and then registering the pseudo-US with the acquired iUS.

Previously proposed methods have largely pursued an improved assessment of image similarity that relies on pre-processed images and/or extended image features. However, they also continue to rely on a global notion of image similarity based on image intensities and other related features throughout the entire image domain. Unfortunately, given the particular image formation model of US, it can be easily argued that most of the image domain that does not involve anatomical boundaries does not provide image information which can truly help in resolving correspondence to an MRI. In that sense, a global assessment of image similarity is simply poor-suited for this particular context and there is instead significant interest in ignoring the US regions which are simply uninformative of underlying anatomical structure. The same issue can be also directly related to the time-sensitive requirements of this particular context. In short, a registration technique that relies on a global assessment of similarity with an extended set of image features can rapidly become impractical in terms of computational cost and processing times. The principle of focusing on exposed anatomical boundaries so as to increase registration robustness and minimize processing times constitutes a critical component of the proposed technique and its benefits are shown in the experimental results section.

The next section enlists a series of experimental findings which show that the proposed approach brings forward gains in computational performance and registration accuracy, as evaluated over fourteen clinical cases obtained from a publicly available dataset. In particular, the proposed technique achieves a robust performance, in the sense that all fourteen cases used for validation have a resulting mean distance between corresponding points that is larger than the smallest possible mean distance (under a rigid transformation) by no more than 1 mm. Furthermore, such performance is achieved with a highly reduced subset of voxels (e.g. 2% of the image) and a GPU-based implementation, which leads to

an average processing time of 0.76 seconds. This achievement permits the technique to be easily embedded in a clinical neuronavigation system, minimizing the delay suffered every time an updated MRI is required.

4.4 Experiments

4.4.1 Clinical Dataset

The validation of the proposed rigid registration technique in the context of IGNS was performed with the use of fourteen clinical neurosurgical cases obtained from the Montreal Neurological Institute's Brain Images of Tumors for Evaluation (MNI BITE) [103], an open access on-line¹ dataset of clinical MRI and US images of brain tumours. In particular, the rigid registration of pre-operative MRI images to iUS was evaluated with images obtained prior to tumour resection, identified as Group 2 of the MNI BITE dataset. The cases involve low and high-grade gliomas (LGG and HGG respectively) at different depths and locations in the brain and with tumour volumes ranging between 0.2 and 79.2 cm³. The initial location of each case corresponds to a preliminary registration involving the manual identification of corresponding points on the skin and the MRI image, as is common in standard clinical procedures.

It is important to highlight that while this intra-operative context does in fact involve non-rigid deformations, a major component of the deformation encountered can be captured by a rigid transformation. Furthermore, Section 4.5 demonstrates that the manual landmarks used for measuring registration accuracy include a significant variability across experts and a relatively small number of locations per case, between 19 and 40 landmarks in total per case, which can critically limit the ability of the landmarks to fully and impartially capture the accuracy of a given deformation across the entire image domain. The experiments in this chapter are thus limited to the validation of a *rigid* registration technique and further work should focus on the validation of non-rigid approaches.

4.4.2 Validation Strategy

Given the lack of a ground truth, the registration accuracy of each method is evaluated as the mean distance between homologous landmarks independently identified by two or

¹<http://www.bic.mni.mcgill.ca/Services/ServicesBITE>

three experts², commonly referred to as the mean target registration error (mTRE). Each case has between 19 and 40 landmarks in total. It is of critical importance to note that the minimal mTRE under a rigid transformation has a unique non-zero value for each case, both for each expert's landmarks as well as for the combined set of all experts' landmarks. There are two main reasons behind this phenomenon. The first is the inherent uncertainty from the experts in accurately identifying anatomical locations in both modalities (particularly in the US volume). Thus, large errors in the identification of landmarks reduce the accuracy of the performance metric and also result in a potentially false large value for the minimal mTRE under a rigid transformation. The second is the potential presence of non-rigid deformations. In particular, in the case of perfectly accurate landmarks, a non-zero minimal mTRE effectively estimates the “residual” part of the deformation that is not fully explained by a rigid transformation. Hence, a large minimal mTRE can reflect the presence of errors in the identified landmarks and/or significant non-rigid components in the true deformation. For the purposes of illustrating the variability of landmark identification between experts and the potential need for a non-rigid registration, a quantitative analysis of the inter-expert variability is presented in Section 4.5.

4.4.3 Results

This section presents a series of quantitative results obtained with the proposed rigid registration technique, referred to as *GOA*, described in Section 3.1.3. It is important to note that such technique involve the maximization of gradient orientation alignment over high gradient magnitude locations selected from the fixed image domain. The voxel selection used in the technique is thus asymmetric, and hence involves the important underlying assumption that selected fixed image boundary locations do indeed have a counterpart in the moving image, while the inverse case of moving image boundaries having a counterpart in the fixed image domain is not necessarily true. This asymmetric correspondence of boundaries across images maps directly to the context involving MRI volumes registered to US volumes. In particular, US volumes generally expose only a limited set of key anatomical boundaries (e.g. tumour boundary, lateral ventricles and falx) which can also be identified in the MRI volume. Consequently, the results shown in this chapter adopt the US volume as the fixed image, whose exposed boundaries have a counterpart in the MRI volume. The

²Note that the third expert in the MNI BITE dataset identified homologous landmarks for the first six cases only.

alternative of choosing the MRI volume as the fixed image volume would likely be ill-fated since many of the anatomical boundaries exposed in the MRI (e.g. grey to white matter interface) simply cannot be identified in the US volume.

The rigid registration approach was evaluated with three simple variants. The first configuration, referred to as *GOA Full Mask*, involves the maximization of gradient orientation alignment of the top 20% locations with the highest gradient magnitude in the reconstructed 3D US volume. The second configuration, referred to as *GOA Subset*, provides reduced processing times and involves the maximization of gradient orientation alignment of 8,000 locations randomly selected from the previously defined top 20% mask. It is important to note that 8,000 locations correspond to approximately 2% of the voxel locations found in the US volume (the exact ratio varies from case to case). The first two configurations were implemented in C++ and run on a computer with an Intel Core 2 Quad Q6700 CPU. The third configuration, *GOA Subset on GPU*, is implemented to run on a NVIDIA GTX 670 video card and was developed to provide highly reduced processing times. It involves the maximization of gradient orientation alignment of 16,000 locations randomly selected from the top 20% mask.

Fig. 4.11 illustrates the images from Case 3 before and after registration with *GOA Subset on GPU*. Notice how key structures like the falx and the lateral ventricles are closely aligned after registration. The registration results for the three configurations and each of the fourteen cases are shown in Fig. 4.12 and Fig. 4.13. In Fig. 4.12, the performance of each configuration is illustrated with relation to a superset of landmarks that includes all experts' landmarks. It is important to note that all configurations consistently improve the registration accuracy with relation to their corresponding initial location. Additionally, one can also observe that the resulting mTRE is very close to the minimal mean distance under a rigid transformation (depicted as a red dashed line). The same set of results is enlisted in Table 4.1, which also shows the processing times for each configuration. The first configuration, *GOA Full Mask*, involves a processing time that ranges from 36 to 76 seconds, which is comparable to the processing time of conventional intensity-based methods. Alternatively, the second configuration, *GOA Subset*, obtains a similar registration accuracy with significantly reduced processing times that range from 7 to 14 seconds. Finally, the third configuration, *GOA Subset on GPU*, also obtains a similar registration accuracy but with highly reduced processing times that range from 0.61 to 0.93 seconds. Note that alternative registration techniques typically involve computational times in the order of minutes. Such

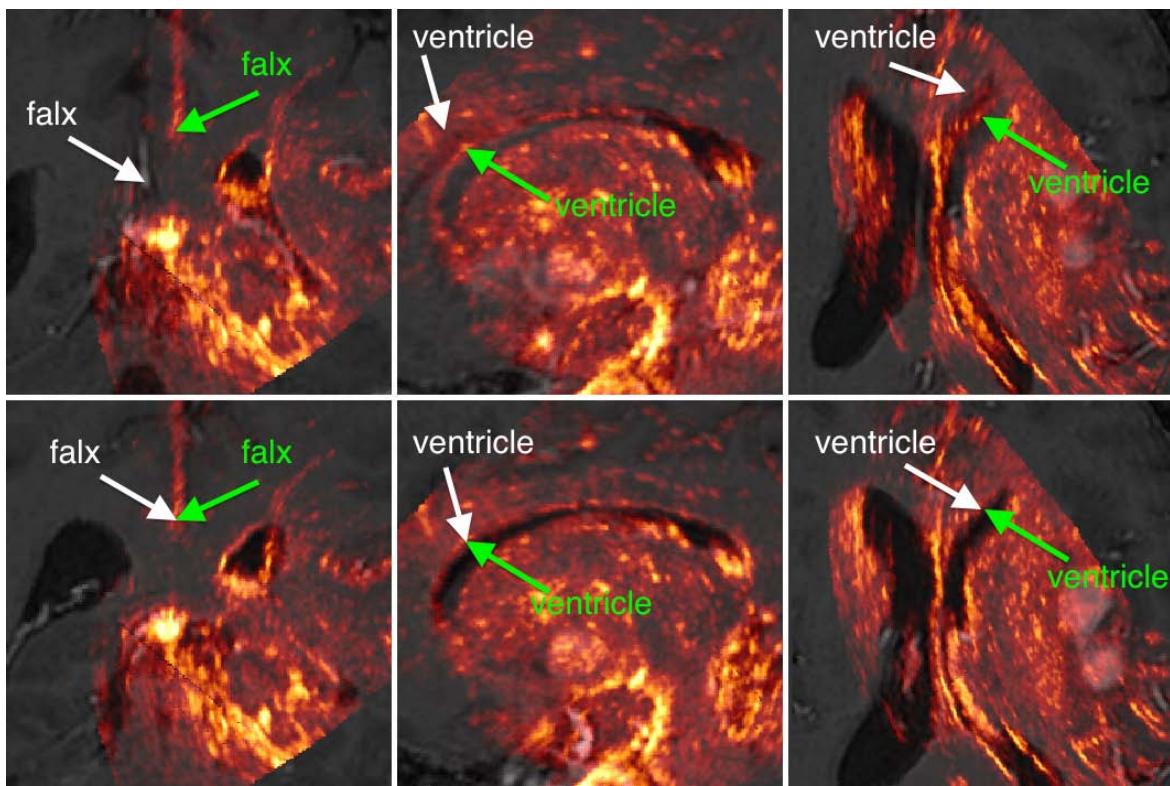


Fig. 4.11 Pre-operative MRI and iUS slices from Case 3 before and after registration. The first row shows the slices at their initial location (coronal, sagittal and transverse). The second row shows the slices after registration with the proposed approach. Anatomical structures found in the iUS are identified by a green arrow and label, while structures found in the MR are identified by a white arrow and label. Notice how key anatomical structures like the falx and the lateral ventricles are tightly aligned after registration with the proposed approach.

a drastic reduction in processing times is critically important in this intra-operative context, since it allows the clinician to immediately obtain updated/corrected pre-operative data.

In Fig. 4.13 the performance of each configuration is illustrated with relation to the set of landmarks identified by each expert. It is immediately clear that the proposed method is in very close agreement with the landmarks selected by Expert 1 and Expert 3, since the resulting mTRE is quite close to their minimal mTRE. On the other hand, when evaluated with relation to Expert 2 landmarks, one encounters a slightly larger value than the minimal mTRE. However, this comes as no surprise, since Section 4.5 will demonstrate that Expert

GOA	Full Mask		Subset		Subset on GPU	
	Case	mTRE in mm	Time in secs	mTRE in mm	Time in secs	mTRE in mm
1	4.89	53	4.86	12	4.86	0.71
2	1.79	50	1.78	9	1.71	0.71
3	2.73	76	2.65	16	2.83	0.93
4	1.68	37	1.72	7	1.76	0.61
5	2.12	36	2.13	9	2.00	0.88
6	1.81	48	1.71	9	1.81	0.68
7	2.51	48	2.64	14	2.75	0.79
8	2.63	56	2.65	10	2.51	0.82
9	2.70	40	2.79	7	2.67	0.86
10	1.95	73	1.94	9	2.05	0.67
11	1.56	52	1.82	9	1.78	0.73
12	2.64	41	2.47	8	2.58	0.72
13	3.47	43	3.42	12	3.57	0.89
14	2.94	62	2.92	9	3.04	0.69
Average	2.53	51	2.54	10	2.57	0.76
Median	2.57	49	2.55	9	2.54	0.73

Table 4.1 Rigid registration results with proposed method evaluated as the mean distance (i.e. mTRE in mm) between all the landmarks independently identified by all experts combined. The results are also illustrated in Fig. 4.12. Also shown is the processing time for each case. Notice that the GPU implementation of the proposed approach yields an average processing time of 0.76 seconds.

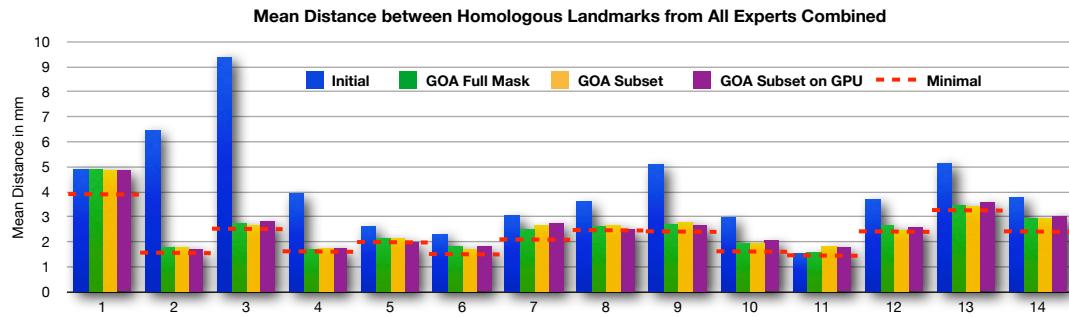


Fig. 4.12 Rigid registration results with proposed method evaluated the set of all landmarks independently identified by three different experts. The x-axis corresponds to the clinical case, while the y-axis corresponds to the mTRE between manually identified corresponding anatomical points. Also shown are the initial mTRE (Initial) in blue, as well as the minimal mTRE (Minimal) possible under a rigid transformation, shown as a red dashed line. Notice that the three configurations yield a mTRE just slightly larger than the minimal mTRE. Table 4.1 also lists the resulting mTRE values for each configuration as well as the corresponding processing times.

2 had significant disagreements with Expert 1 and 3. Thus, any method that tends to align more with landmarks from Expert 1 and 3, will inevitably show a degradation with relation to the landmarks from Expert 2.

A statistical summary of the registration performance of the proposed approaches as well as of competing techniques is presented in Table 4.2. The evaluated methods are characterized by a choice of similarity metric: Mutual Information (MI), Normalized Cross Correlation (NCC), Normalized Mutual Information (NMI) and Gradient Orientation Alignment (GOA); input images: Original Images (ORI), Gradient Magnitude Images (GM) and Median Filtered MRI in conjunction with a Gaussian blurred US (PRE); and optimization strategy: Covariance-Matrix-Adaptation-Evolution-Strategy (CAE) and a Gradient Descent (GD) optimizer with adaptive gain. The mean and median value of two accuracy measures are reported: the mTRE of the homologous landmarks from all experts combined, and the difference between such mTRE and the minimal mTRE. The number of cases that had a successful registration are also demonstrated, where success is defined as an instance when the mTRE is larger than the minimal mTRE by no more than 1 or 2 mm.

The results obtained indicate that conventional multi-modal intensity-based metrics, like MI, NCC and NMI, generally show very poor performance in this particular context.

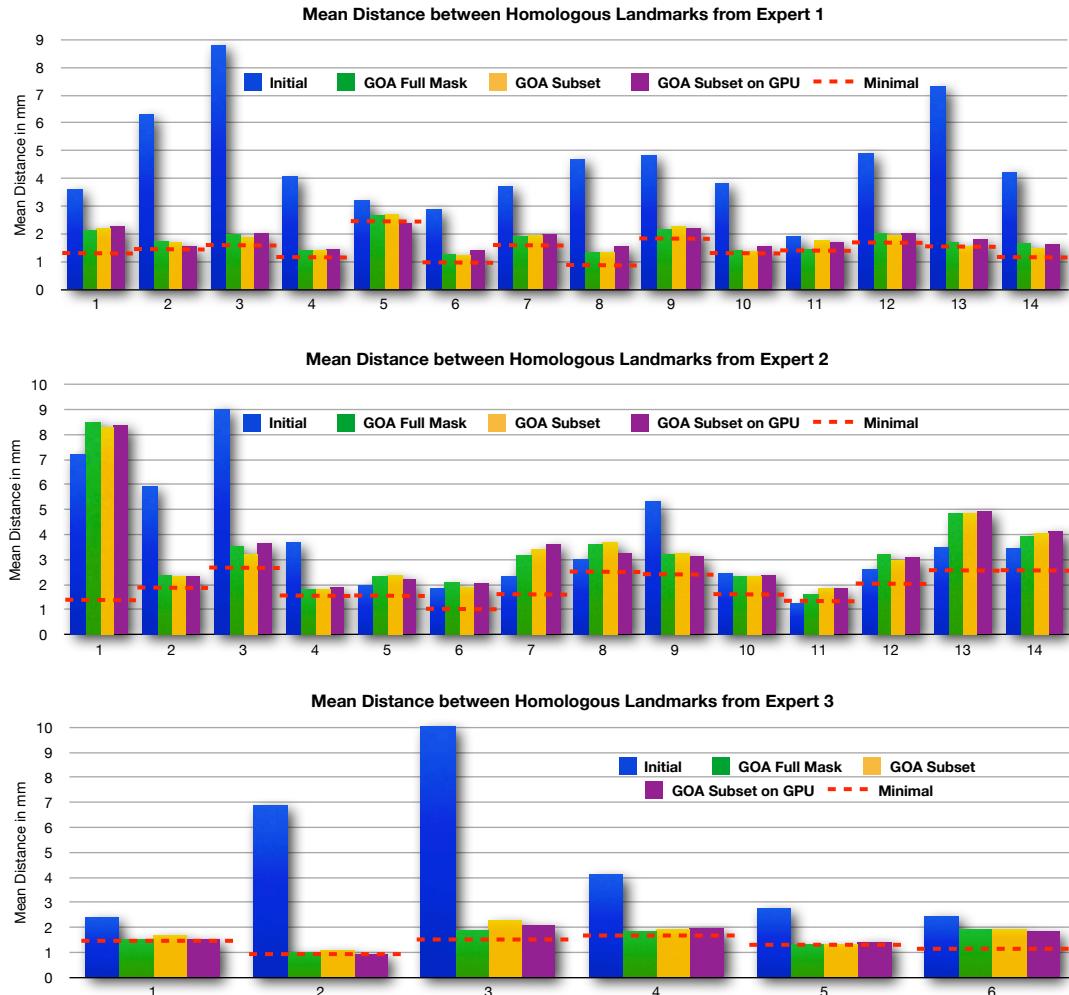


Fig. 4.13 Rigid registration results with proposed method evaluated with each of three sets of landmarks identified by a particular expert. The x-axis corresponds to the clinical case, while the y-axis corresponds to the mTRE between manually identified corresponding anatomical points. Also shown are the initial mTRE, (Initial), in blue, as well as the minimal mTRE (Minimal) possible under a rigid transformation, shown as a red dashed line. Notice that for Expert 1 and 3, all three configurations yield a mTRE just slightly larger than the minimal mTRE and significantly decreased with relation to the initial mTRE. However, evaluation with relation to Expert 2 yields various cases with poor performance (e.g. Case 1, 7, 8, 12, 13 and 14).

Nonetheless, if one makes use of gradient magnitude images with the same multi-modal similarity metrics, one can obtain slightly improved results. A particularly good-performing configuration involves the maximization of NMI between gradient magnitude images with a gradient-descent strategy, which successfully registers 12 of the 14 cases with an mTRE less than 1 mm larger than the minimal mTRE.

In comparison with other competing techniques evaluated, the proposed approach is the only one where all cases were successfully registered to an mTRE that was less than 1 mm larger than the minimal mTRE. In particular, the median value of the difference between the mTRE and the minimal mTRE is of 0.27 mm for the first configuration, (GOA Full Mask); 0.22 mm for the second configuration, (GOA Subset); and 0.33 mm for the third configuration, (GOA Subset on GPU).

Method	mTRE in mm		mTRE minus minimal mTRE in mm		Number of Successes	
	Average	Median	Average	Median	<1mm	<2mm
MI+ORI+CAE	22.62	9.20	20.42	7.51	3	3
NMI+ORI+CAE	21.58	8.63	19.39	6.89	3	4
NCC+ORI+CAE	70.58	76.30	68.38	74.11	0	0
MI+ORI+GD	11.83	6.84	9.64	4.44	3	5
NMI+ORI+GD	11.68	6.88	9.64	4.44	3	5
NCC+ORI+GD	37.83	29.73	35.63	28.21	0	0
MI+GM+CAE	3.00	2.90	0.81	0.64	9	12
NMI+GM+CAE	3.01	2.78	0.82	0.59	10	12
NCC+GM+CAE	6.59	3.09	4.39	0.79	8	11
MI+GM+GD	2.90	2.67	0.71	0.61	11	14
NMI+GM+GD	2.87	2.66	0.67	0.60	12	13
NCC+GM+GD	3.77	2.91	1.57	0.69	9	12
NMI+PRE+CAE	13.47	9.22	11.28	6.83	0	1
NMI+PRE+GD	10.78	8.67	8.59	5.83	0	0
GOA Full Mask	2.53	2.57	0.33	0.27	14	14
GOA Subset	2.54	2.55	0.34	0.22	14	14
GOA Subset on GPU	2.57	2.54	0.37	0.33	14	14

Table 4.2 Statistical summary of rigid registration results with all evaluated techniques. The first column identifies the method used which are characterized by a similarity metric (e.g. Mutual Information (MI), Normalized Cross Correlation (NCC), Normalized Mutual Information (NMI) and Gradient Orientation Alignment (GOA)), input images (e.g. Original Images (ORI), Gradient Magnitude Images (GM) and Median Filtered MRI in conjunction with a Gaussian blurred US (PRE)). and optimization strategy (e.g. Covariance-Matrix-Adaptation-Evolution-Strategy (CAE) and a Gradient Descent (GD) optimizer with adaptive gain). The first two columns of registration results show the average and median value of the mTRE between manually identified corresponding points. Also shown are the average and median value of the mean distance minus the minimal mTRE. The last columns indicate the number of cases in which the resulting mTRE is larger than the minimal mTRE by no more than 1 or 2 mm respectively.

4.5 Inter-Expert Variability

The validation of image registration is a general issue across many medical imaging contexts, and it is particularly important and challenging in the context of IGNS. The importance lies in the fact that the validation strategy will ultimately characterize the precision of a registration technique and provide the clinician with a measure of how much to “trust” the updated pre-operative imaging data. Hence, a poorly designed validation strategy may wrongly evaluate a registration technique as having a high accuracy and can consequently lead to a mistaken decision of whether a particular region of tissue should be resected or not. The particular challenge lies in the fact that iUS is an image modality that is relatively hard to interpret by experts, and there is simply no easy approach for accurately identifying a dense set of homologous correspondences. This section provides an analysis of the agreement between the three experts which provided homologous landmarks so as to evaluate the performance of a registration algorithm. In short, the objective is to illustrate the challenges of the validation strategy with a series of quantitative results.

For each expert’s landmarks, the initial mTRE of each case is evaluated. The mTRE evaluated with the optimal rigid transformation is also reported. The optimal rigid transformation given a set of landmarks is obtained analytically by solving the corresponding Orthogonal Procrustes problem [149]. Note that the mTRE evaluated with the optimal rigid transformation corresponds to the minimal mTRE that can be obtained with a rigid transformation given that set of landmark points. In order to analyze the agreement between different experts, the mTRE of a given expert’s landmarks with relation to the optimal rigid transformation obtained with *another expert’s* landmarks is also evaluated. Such measure is particularly informative in cases that do not appear to require a non-rigid transformation.

Tables 4.3-4.5 list the results from such analysis. For example, Table 4.3 evaluates the mTRE based on the landmarks identified by Expert 1. The column labeled *Initial* lists the mTRE values computed at the initial (unregistered) location of each case. The column labeled *Expert 1 Solution* lists the mTRE values computed with the optimal rigid transformation (analytically) obtained from Expert 1’s landmarks. Hence, the column labeled *Expert 1 Solution* in Table 4.3 also lists the minimal possible mTRE. All values of the minimal mTRE larger than 2 mm (highlighted in bold) are identified as cases with potential need for a non-rigid model. Notice that between Expert 1 and 3, only Expert

mTRE based on Expert 1 Landmarks		Expert 1 Solution	Expert 2 Solution	Expert 3 Solution
Case	Initial	(Minimal)		
1	3.62	1.23	11.55	2.11
2	6.29	1.38	1.86	1.47
3	8.79	1.63	1.98	2.25
4	4.07	1.10	1.58	1.48
5	3.20	2.37	2.51	2.47
6	2.89	0.96	2.11	1.93
7	3.70	1.71	2.52	-
8	4.70	0.89	2.86	-
9	4.82	1.89	2.71	-
10	3.83	1.26	2.15	-
11	1.90	1.39	1.72	-
12	4.89	1.67	3.84	-
13	7.30	1.45	4.79	-
14	4.23	1.09	2.58	-
Average	4.59	1.43	3.20	1.95
Median	4.15	1.38	2.52	2.02

Table 4.3 Analysis of homologous landmarks identified by Expert 1. Reported numbers include the initial mTRE (Initial), and the minimal mTRE, found with the analytical rigid transformation obtained from each expert's landmarks. Minimal mTRE values larger than 2 mm are highlighted in bold and expose a potential need for a non-rigid registration. Values of mTRE larger than 3 mm are shown in red bold and indicate a notable difference between experts.

1 yields one case with a minimal mTRE larger than 2 mm. In direct contrast, Expert 2 yields six cases with a minimal mTRE larger than 2 mm. Hence, one could argue that for most of the evaluated cases and their corresponding landmarks, a rigid transformation can characterize most of the deformation encountered.

In Tables 4.3-4.5, all mTRE values larger than 3 mm and evaluated with relation to a different expert's solution are identified as cases with significant disagreement between experts (highlighted in red bold). The most striking case of disagreement is found in Case 1, where each expert has a corresponding minimal mTRE of 1.23, 1.23, and 1.35 mm. Thus, there seems to be no strong need for a non-rigid transformation. However, the rigid solution obtained with the landmarks from Expert 2 results in a mTRE value of 11.55 mm with

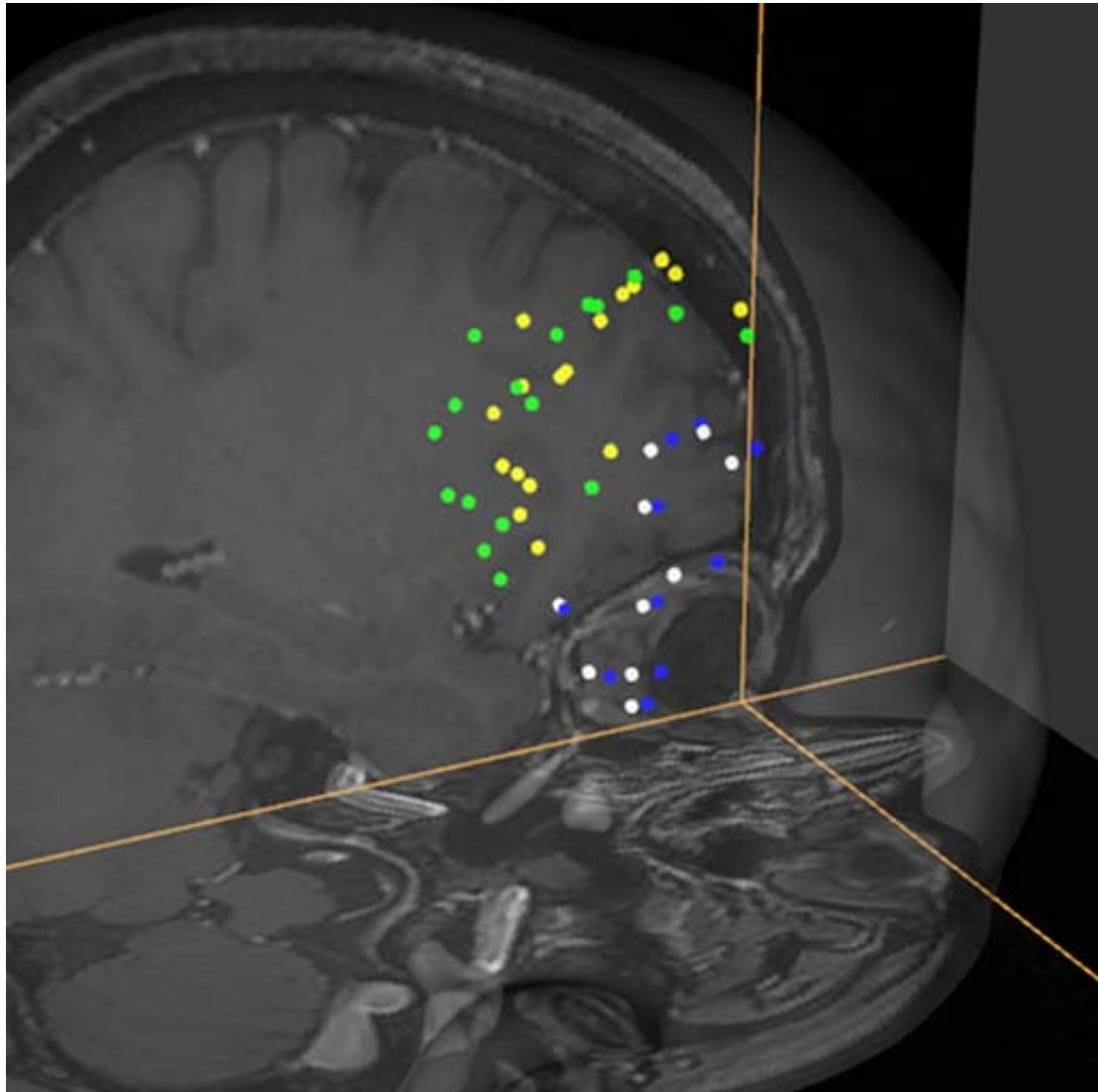


Fig. 4.14 Homologous landmarks identified by Expert 1 and 2 for Case 1. Points identified by Expert 1 are colored in blue (MR) and white (US). Points identified by Expert 2 are colored in green (MR) and yellow (US). A coronal and transverse slice of the MRI are also shown for reference, as well as a translucent rendering of the skin's surface.

mTRE based on Expert 2 Landmarks		Expert 1 Solution	Expert 2 Solution	Expert 3 Solution
Case	Initial	(Minimal)		
1	7.20	10.28	1.23	8.11
2	5.93	2.29	1.91	2.27
3	9.02	3.10	2.86	3.92
4	3.68	2.23	1.61	1.88
5	1.96	2.12	1.67	2.21
6	1.86	2.45	1.04	2.16
7	2.31	3.10	1.74	-
8	3.02	3.83	2.49	-
9	5.33	3.36	2.39	-
10	2.43	2.48	1.71	-
11	1.26	1.53	1.22	-
12	2.61	3.66	2.04	-
13	3.49	5.40	2.56	-
14	3.45	3.62	2.57	-
Average	3.82	3.53	1.93	3.42
Median	3.23	3.10	1.83	2.24

Table 4.4 Analysis of homologous landmarks identified by Expert 2. Reported numbers include the initial mTRE (Initial), and the minimal mTRE, found with the analytical rigid transformation obtained from each expert’s landmarks. Minimal mTRE values larger than 2 mm are highlighted in bold and expose a potential need for a non-rigid registration. Values of mTRE larger than 3 mm are shown in red bold and indicate a notable difference between experts.

Expert 1’s landmarks. In contrast, the rigid solution obtained with the landmarks from Expert 3 results in a mTRE value of 2.11 mm with Expert 1’s landmarks. Additionally, the rigid transformation from Expert 2 results in a mTRE of 10.69 mm when evaluated with the landmarks chosen by Expert 3. In contrast, the rigid transformation from Expert 1 results in a mTRE of 2.66 mm when evaluated with Expert 3’s landmarks. Hence, one can argue that Expert 1 and 3 seem to somewhat agree on the deformation encountered in Case 1, while having a strong disagreement with Expert 2.

The landmarks identified by Expert 1 and 2 for Case 1 are illustrated in Fig. 4.14 for further analysis. Notice the significant difference between the spatial distribution of each set, and the difference between the apparent transformation for each set. In particular, the

mTRE based on Expert 3 Landmarks		Expert 1 Solution	Expert 2 Solution	Expert 3 Solution
Case	Initial	(Minimal)		
1	2.39	2.66	10.69	1.35
2	6.87	1.10	1.47	0.88
3	10.69	2.49	3.06	1.50
4	4.10	2.11	1.87	1.70
5	2.76	1.79	1.86	1.30
6	2.45	2.16	1.95	1.10
Average	4.88	2.05	3.48	1.31
Median	3.43	2.13	1.91	1.33

Table 4.5 Analysis of homologous landmarks identified by Expert 3. Reported numbers include the initial mTRE (Initial), and the minimal mTRE, found with the analytical rigid transformation obtained from each expert’s landmarks. Minimal mTRE values larger than 2 mm are highlighted in bold and expose a potential need for a non-rigid registration. Values of mTRE larger than 3 mm are shown in red bold and indicate a notable difference between experts. Note that Expert 3 only labelled the first six of fourteen cases in the dataset.

points identified by Expert 1 (shown in blue and white) are already quite close in distance, while the ones identified by Expert 2 (shown in green and yellow) are significantly farther apart. For reference purposes, Fig. 4.15 illustrates the landmarks identified by Expert 1 and 2 for Case 13, which exposes a less prominent disagreement between experts. In such case, the distributions of the two landmark sets are relatively similar when compared to the ones found in Case 1.

The relevance of the inter-expert analysis on the landmarks lies in highlighting the challenges involved in the validation of a registration method in real clinical cases with no ground truth. In particular, it is important to underline that although the use of manually identified points allows for a quantitative evaluation of performance, there is still a significant degree of subjectivity behind such validation strategy and the numerical results should not be accepted blindly. In particular, based on the exposed variability between experts and corresponding landmarks, one can argue that the validation of a non-rigid registration based on this particular dataset is rather compromised.

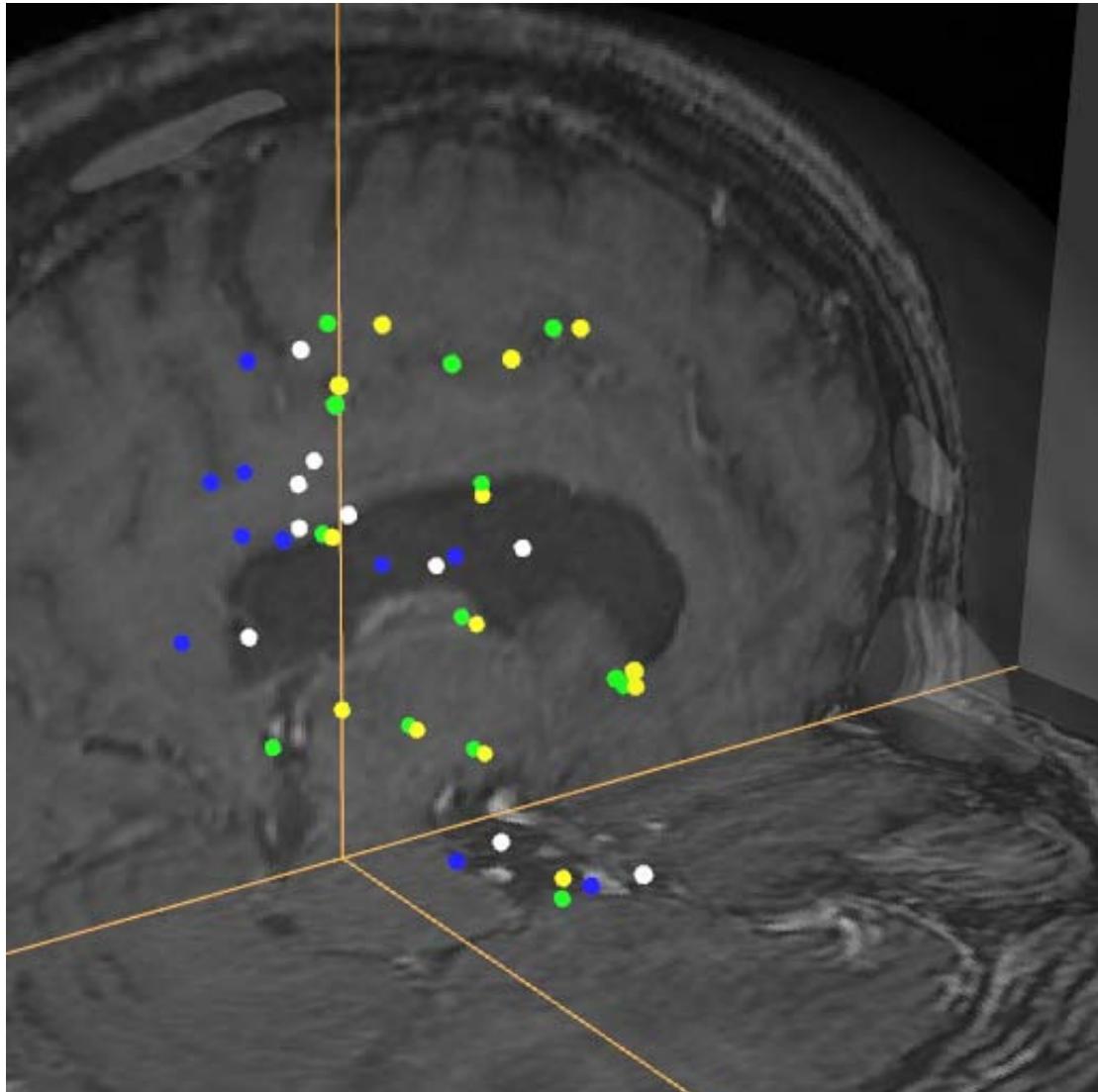


Fig. 4.15 Homologous landmarks identified by Expert 1 and 2 for Case 13. Points identified by Expert 1 are colored in blue (MR) and white (US). Points identified by Expert 2 are colored in green (MR) and yellow (US). A coronal and transverse slice of the MRI are also shown for reference, as well as a translucent rendering of the skin's surface.

4.6 Impact

4.6.1 IBIS NeuroNav

The proposed registration techniques evaluated in this section have been integrated into a prototype neuronavigation system, referred to as IBIS NeuroNav[104], which has been custom developed at the Montreal Neurological Institute. IBIS provides a wide range of functionality for neuronavigation including: loading of medical volumes and segmented surfaces, 3D visualization, integration with 3D tracking technology, support for capture and visualization of tracked ultrasound, among many others.

A key feature of IBIS is that it includes a modular plug-in architecture that allows developers to extend its functionality. Such architecture was leveraged so as to extend IBIS with robust registration techniques based on the proposed strategy in conjunction with a fast 3D reconstruction implementation which jointly greatly reduce the overhead of using iUS in the operating room. The following section briefly outlines the technical contributions developed as an extension of IBIS.

Technical Contributions to IBIS

GPU-based reconstruction of a 3D US volume To minimize the computational overhead of integrating tracked US in the operating room, a GPU-based 3D reconstruction algorithm was implemented as a plug-in for IBIS. The 3D reconstruction algorithm takes as input a series of masked 2D US images with known location and orientation, and constructs a 3D volume based on the interpolated values of the 2D US slices. The GPU-based reconstruction algorithm involves processing times that typically range between 5 to 10 seconds.

In it important to note that the prior protocol involved the used of a Perl script, referred to as *volregrid*, which is part of the MINC tool distribution (<http://packages.bic.mni.mcgill.ca>). The script had to be executed as a standalone (independently of IBIS) application and involved processing times in the order of minutes. Once the script had finished, the reconstructed 3D volume had to be manually loaded into IBIS.

The new protocol based on the GPU-based reconstruction allows for a much more seamless integration of iUS in the operating room, since the processing times are significantly reduced and the reconstructed volume is automatically loaded and visualized within IBIS.

CPU-based and GPU-based multi-modal rigid registration Two rigid registration implementations, based on the framework in Section 3.1, were implemented in both a CPU and in a GPU for further accelerated computational times. The GPU-accelerated implementation involved some non-trivial modifications and was specifically developed so as to efficiently exploit the massively parallel nature of both the algorithm and the hardware. The details regarding the GPU-based registration technique can be found in Section 4.6.1.

IBIS has been implemented in C++ and relies on the Insight Toolkit Library. As such, it was straightforward to port the standalone registration code used for research experiments into an IBIS plugin. Nonetheless, there was a special effort required so as to ensure that the transformation hierarchy of all objects in the 3D scene in IBIS was accounted for by the registration plugin. For example, in the case of neurosurgery the 3D scene typically has a set of 3D volumes and segmented surfaces grouped under a “Pre-operative” group. The images obtained intra-operatively are separately grouped under an “Intra-operative” group. Thus, when rigidly registering a pre-operative MRI to an iUS it is critically important to apply the transformation obtained to the entire “Pre-operative” group, so that all pre-operative data can be properly updated.

Accelerated Implementation on Graphical Processing Unit

This section briefly describes the modifications that had to be done on the proposed registration technique so as to maximize the computational throughput of the graphical processing unit. Note that GPU-specific development was performed in OpenCL.

The sole distinction between the CPU-based implementation and the GPU-based implementation lies in the evaluation of the similarity function, which largely represents the most computationally expensive component of the registration technique. All other components, such as the initial computation of the image gradients, the selection of a reduced subset of locations of interest, and the overhead operations performed by the optimizer, are performed in the CPU with a common C++ implementation.

The evaluation of the local similarity function can be rendered highly efficient by carefully considering the memory access limitations of the GPU. In particular, a common challenge in maximizing the computational efficiency of a technique in a GPU lies in maximizing the data throughput between the GPU memory bank and the parallel processing units. Fig. 4.16 illustrates how the data is allocated in the GPU in the proposed imple-

mentation for the purposes of evaluating the similarity function efficiently in parallel. The strategy for achieving parallelized memory access is known as *memory coalescing*, where the objective is to ensure that parallel processing units jointly load the input data, and thus require one single memory read request. An important requirement for achieving this goal is that the input data used by the processing units should be 16-byte aligned and ordered in the same order as the processing units are ordered. When the algorithm is setup in such a fashion, the GPU can load the complete input data array in a single read request to the memory buffer.

In the proposed technique, the input array is composed of data elements each involving 8 single floating-point values. In particular, the data structure stores a fixed image gradient, $\nabla I_f(\mathbf{x})$, and its corresponding location, \mathbf{x} . To ensure memory coalescing, two unused floating-point values are also included, such that the parallel memory read request remains 16-byte aligned. The original moving image gradient is stored in a four-component 3D image in the GPU memory. Note that similar to the input array, voxel values are 16-byte aligned so as to maximize memory efficiency.

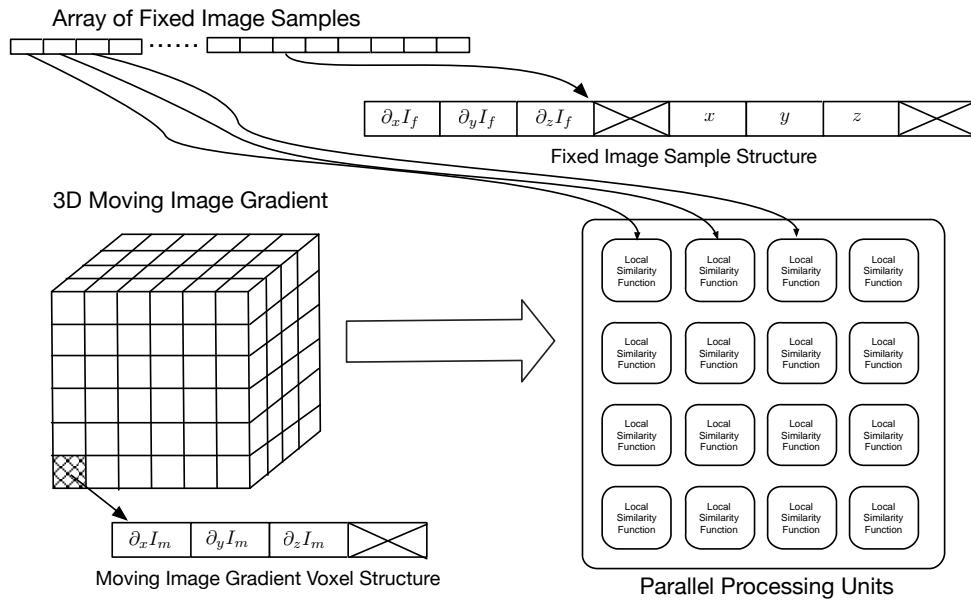


Fig. 4.16 GPU data allocation for evaluating the similarity function in parallel.

The evaluation of the similarity function requires the corresponding transformed moving

image gradient, ∇I_m^\downarrow . In order to evaluate the transformed moving image gradient, as expressed in Equation 3.13, one first requires the transformed fixed image location, $\mathbf{T}(\mathbf{x})$, so that one can evaluate the interpolated value of the original (undeformed) moving image gradient. In the context of rigid registrations, one can compactly express the relationship between a given fixed image location, $\mathbf{x} = (x, y, z)$, and the spatially mapped moving image index, (i, j, k) , with a single homogeneous matrix such as,

$$\begin{pmatrix} i \\ j \\ k \\ 1 \end{pmatrix} = \mathbf{M}_{f \rightarrow m} \cdot \mathbf{T}_h \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (4.1)$$

$$= \mathbf{T}_{f \rightarrow m} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (4.2)$$

where \mathbf{T}_h is the rigid transformation matrix expressed in homogeneous coordinates and where $\mathbf{M}_{f \rightarrow m}$ is a matrix which maps a location in the moving image spatial domain to a continuous voxel index. Hence, one can join both matrices into a single matrix, $\mathbf{T}_{f \rightarrow m}$, which maps a fixed image location to a continuous voxel index in the moving image. Finally, the interpolated moving image gradient value is obtained by relying on hardware-implemented image interpolation units found in the GPU.

The interpolated moving image gradient is then pre-multiplied by the rotation matrix, $\mathbf{R} = \mathbf{J}^T$, to obtain the transformed moving image gradient. One can then compute the value of the local similarity metric. The algorithmic outline of the sequence performed by each parallel processing unit is summarized in Algorithm 3.

4.7 Conclusions

Brain tumour resection is an aggressive option for improving the prognosis of suffering patients, yet in some cases it is the only alternative towards preventing an imminent death. Studies have shown that the extent of surgical resection is an important factor for the prognosis of the patient and the use of intra-operative imaging solutions, such as US, is an

Algorithm 3 Local similarity function in parallel processing unit.

```

1: procedure SIMILARITY( $id, \mathbf{T}_{f \rightarrow m}, \mathbf{R}$ )
2:    $\nabla I_f(\mathbf{x}), \mathbf{x} \leftarrow \text{loadData}(id)$ 
3:
4:   
$$\begin{pmatrix} i \\ j \\ k \\ 1 \end{pmatrix} \leftarrow \mathbf{T}_{f \rightarrow m} \times \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

5:
6:    $\nabla I_m(\mathbf{T}(\mathbf{x})) \leftarrow \text{interpolate}(\nabla I_m, [i, j, k])$ 
7:
8:    $\nabla I_m^\downarrow(\mathbf{x}) \leftarrow \mathbf{R} \times \nabla I_m(\mathbf{T}(\mathbf{x}))$ 
9:
10:  return  $s(\nabla I_f(\mathbf{x}), \nabla I_m^\downarrow(\mathbf{x}))$ 
11: end procedure

```

appealing technical strategy for optimizing tumour resection while minimizing the damage on eloquent tissue.

In this chapter, I have described the technical setup involving the use of freehand tracked iUS for improving the accuracy of IGNS. I have enlisted two key technical contributions that extend the functionality of IBIS Neuronav and that allow a more seamless integration of iUS in the operating room. The first contribution involves the implementation of a GPU-based volume reconstruction plug-in, which allows the clinician to obtain a 3D US volume immediately after an acquisition sweep. The second contribution involves the implementation of a CPU and a GPU-based MRI to iUS rigid registration technique for correcting image misalignments suffered during surgery.

The performance of the proposed registration technique was evaluated on a publicly available dataset of real clinical cases. A key result of such set of experiments is that the GPU-based technique can provide a highly robust rigid registration performance with sub-second processing times. Such performance in conjunction with a GPU-based volume reconstruction technique provides an extremely fast turnaround time for the clinician and thus minimizes the overhead of adopting iUS.

The next chapter evaluates the performance of the proposed rigid registration technique in a multi-modal context involving brain CT and MRI volumes with the use of a publicly available dataset. In particular, it presents a series of experimental results that illustrate the performance of the proposed technique with varying parametrizations and compares

the performance against various previously proposed approaches.

Chapter 5

Rigid Registration of CT to MRI Brain Volumes

The previous chapter described in detail the clinical relevance and impact of the proposed registration technique in the challenging context of IGNS based on iUS. This chapter focuses on the multi-modal rigid registration of CT and MRI brain volumes, which constitutes a more conventional context that has been previously approached by many different groups [30, 35, 57, 119, 172]. Registration performance is evaluated with a commonly used and publicly available dataset [172] and thus allows for a direct and fair comparison with previously reported results. This chapter also provides a detailed quantitative analysis of the performance of the proposed technique with a variety of different parameterizations. Since such analysis is also based on the same publicly available dataset [172] it provides an unbiased representation of how the method's parameters affect its accuracy.

The chapter is structured as follows. Section 5.1 describes the variety of clinical contexts that necessitate the rigid registration of brain CT and MRI volumes. Section 5.2 provides a brief overview of the publicly available dataset used for evaluation of registration performance as well as describing its corresponding validation strategy. Section 5.3 presents a brief recapitulation of the rigid registration technique and the specific parameters involved. Section 5.4 then presents a variety of quantitative experimental results. The first set of results characterize the performance of the proposed technique with varying parameterizations, while the second set of results focuses on comparing registration performance with previously proposed techniques. The comparison against competing techniques includes a

previously proposed registration strategy [119] which similarly relies on image gradients as critical image features of interest, and thus serves as an important reference regarding previous efforts that follow such intuition. Nonetheless, there are important distinctions in the way image gradient information is employed in such method and the proposed method, and they are discussed with the corresponding experimental results.

The results presented in this chapter provide significant evidence of the generality of the proposed technique to various multi-modal contexts, as well as of the important benefits that can be obtained in terms of robustness, accuracy, and computational efficiency. For example, the performance comparison includes an evaluation of registration accuracy as a function of the percentage of voxels used to estimate the similarity metric, where the proposed method achieves a median registration error inferior to 1 mm with an extremely reduced sampling rate involving 0.05% of the voxels in the domain. In contrast, a state-of-the-art pixel selection technique achieved a comparable yet slightly worse registration accuracy with 5% of the voxels (i.e. 100 times more voxels).

5.1 Clinical Applications

The need for rigid-registration of multi-modal brain volumes commonly arises in the context of surgical guidance and planning where the clinician may desire to jointly analyze a series of brain images from different modalities in a common coordinate frame. Since the image coordinate space of each modality generally is arbitrary and not matched across modalities, there is an immediate need to register all acquired images into a common coordinate space. A similar issue arises in longitudinal studies where a patient is scanned at multiple time points and the clinician is interested in identifying any anatomical changes that occurred through time. For example, a clinician may rely on a series of MRI modalities across multiple time points so as to identify the evolution of a brain tumour. In such context, accurately registered multi-modal images across time points allow for a localized (e.g. per pixel) analysis of changes through time.

5.2 Retrospective Image Registration Evaluation Project

The experiments shown in this chapter are evaluated with real clinical cases obtained from the Retrospective Image Registration Evaluation (RIRE) Project [172], also known as the

Vanderbilt dataset. The RIRE dataset is an open-access dataset of rigid registration cases involving CT, MRI and PET brain volumes. The MRI volumes found in the dataset include T1-weighted, T2-weighted and PD-weighted modalities. Furthermore, some cases also include rectified versions of MRI modalities, which correct MRI volumes for static field inhomogeneity by using an image rectification technique [27].

The RIRE project also provides a “blind” validation tool in their website, in which any researcher can upload a set of transformations obtained by their algorithm and the website will provide a quantitative description of the registration accuracy obtained. The validation strategy is “blind” because the user does not know the gold standard used for evaluating accuracy. Furthermore, all uploaded results are stored in the website and can be verified by other research groups. As such, the RIRE dataset provides a very controlled strategy for comparing registration performance across competing methods.

5.3 Methodology

The rigid registration technique evaluated in this section is described in detailed in Section 3.1.1 and it involves the maximization of a local similarity metric evaluating gradient orientation alignment of a highly reduced set of selected voxel locations chosen from within a top set of high gradient magnitude voxels. In particular, the metric is expressed as,

$$S_l(\nabla I_f(\mathbf{x}), \nabla I_m^\downarrow(\mathbf{x})); N = \cos(\theta(\mathbf{x}))^N \quad (5.1)$$

where $\theta = \angle(\nabla I_f, \nabla I_m^\downarrow)$ is the inner angle between the fixed image gradient, ∇I_f , and the transformed moving image gradient, ∇I_m^\downarrow ; and N is a parameter characterizing the selectivity with which to evaluate gradient orientation alignment.

The proposed technique is *asymmetric* in the sense that the fixed image plays a different role than the moving image. In particular, the locations where the local similarity metric is evaluated are selected from the fixed image domain. As such, it implicitly assumes that selected locations have an anatomically corresponding image boundary in the moving image. In the results shown in this chapter, the fixed image modality is chosen to be CT, since it strongly and accurately depicts bone-tissue and skin-tissue boundaries which can be also found in MRI volumes. The alternative choice of choosing the MRI volumes as the fixed image modality could likely lead to decreased robustness since most of the soft-tissue

(e.g. gray matter, white matter) boundaries found in MRI volumes are simply not exposed in the CT volume.

5.4 Results

In this section, the performance of the proposed registration technique is evaluated with relation to varying parameters that include: the matching criterion, N , the uncertainty (i.e. gradient magnitude) threshold, the sampling rate and the use of (or lack of) a sub-voxel location sampling. The different configurations of the method are evaluated with the use of the training set found in the RIRE dataset. The training set involves the registration between one CT volume and six MRI volumes (three modalities with and without rectification) and provides the ground truth transformation for each registration task. For this particular set of experiments, the registration performance of each case is evaluated as the mean distance between the estimated and gold truth location of the eight vertex points defining the CT volume boundary. Such a metric can be considered an approximation of the upper bound on the mean target registration error (mTRE) within the full CT volume, and will be referred to as the mean vertex registration error (mVRE). The performance across all training cases is then evaluated as the median value of the mVRE. One should note that the performance metric used for the training set is not the same as the one used of the testing set, which is blindly evaluated through the dataset's website.

5.4.1 Random Sub-Voxel Location Sampling

Image volumes in the RIRE database are largely anisotropic and significantly coarse in the out-of-plane dimension (e.g. one volume in the dataset has a voxel spacing of $1.27 \times 1.27 \times 4.11$ mm). Such a largely anisotropic voxel spacing can lead to a potential registration degradation particularly with relation to high-curvature boundaries (e.g. corners) where the gradient orientation has a significant variation within the voxel volume. This constitutes valuable image information content for the proposed technique. If one chooses a sampling strategy limited to on-grid locations one can potentially miss such variation in the gradient orientation. In order to avoid this kind of degradation, a scheme that allows the evaluation of the local metric at locations that do not exclusively fall in the image grid is

introduced. In particular, once a voxel of interest is selected, an alternative sub-voxel location within the voxel's volume is identified by adding a random offset to the voxel location. Note that once one chooses a location that does not fall in the voxel grid, one is obliged to also adopt an interpolation strategy to compute the fixed image gradient orientation. The influence of incorporating such random offset is evaluated with different values of the matching criterion, N , for a fixed sampling mask defined by the top 10% of highest gradient magnitude voxels. Fig. 5.1 illustrates the median value of the mVRE obtained with and without random offsets. The comparison between using and not using a random offset for evaluated locations consistently demonstrates that the offset brings forward a significant improvement in registration accuracy. Thus, all subsequent experiments will include such sub-voxel location sampling scheme.

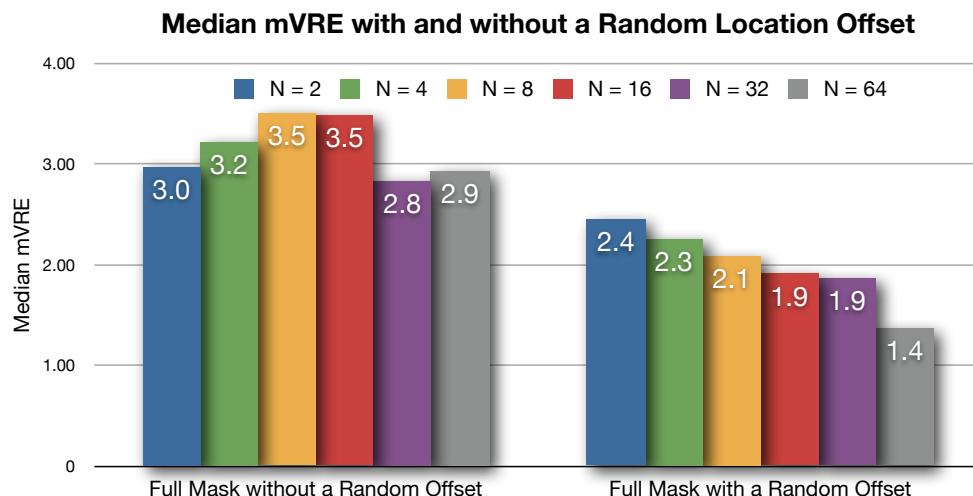


Fig. 5.1 Median mVRE using all voxels within a sampling mask defined by the top 10% voxels with highest gradient magnitude. The plot on the left evaluates the similarity measure on locations specified by the voxel grid, while the plot on the right embeds a random offset within the voxel's volume. Additionally, multiple values of the matching criterion, N , see Equation 5.1, are also evaluated.

5.4.2 Sampling Rate

In the next set of experiments, registration performance is evaluated as a function of the number of locations used (i.e. sampling rate). Since the subset of pixels used for registration

is randomly selected and fixed through the optimization, performance is evaluated across multiple realizations of each registration task. In particular, each individual registration case is evaluated with twenty random instances. Fig. 5.2 illustrates the median mVRE with a varying number of pixels. The average processing time for a given number of pixels is also shown. For reference, the registration performance obtained using all pixels within the sampling mask is also shown. Similar to previous experiments, one can notice that the registration accuracy tends to be higher for large values of N , see Equation 5.1. In particular, one can observe that the smallest value of N , consistently yields the worst performance for all configurations. The “optimal” value of N is not consistent across different numbers of evaluated voxels. Nonetheless, when using a subset of voxels, one can observe that a value of N of 16 or 32 tends to provide optimal or close to optimal results. It is important to highlight that one can resolve most of the transformation with very few voxels and in a short amount of time. In particular, with one thousand voxels, namely 0.013% of all voxels in the volume, one can register the volumes with a median mVRE of 3.1 mm and with an average processing time of 6 seconds (of which 3 seconds are devoted to computing image gradients and the sampling mask).

5.4.3 Sensitivity to Magnitude Threshold

The last component explored is the gradient magnitude threshold defining the sampling mask. For this set of experiments, the value of N is fixed to 32 and the method continues to incorporate a random offset for sub-voxel location sampling. Five masks evaluated with varying thresholds are illustrated in Fig. 5.3. The first and most restrictive mask is defined by the top 1% high gradient magnitude voxels, while the last and least restrictive mask is defined by the top 40% high gradient magnitude voxels. The registration performance is also evaluated with a reduced number of selected voxels. In particular, the method randomly selects a reduced number of voxels (e.g. 8,000) within a particular mask and restricts the optimization of the metric to those fixed locations. Fig. 5.4 illustrates the registration performance evaluated with different masks. The sampling mask defined by the top 10% of the voxels shows consistently better performance when evaluating a randomly selected subset of locations. Furthermore, when using the full sampling mask we observe a relatively similar performance for the masks defined by the top 10%, 20% or 40% of the voxels. In other words, once voxels found in the top 10% mask have been selected, the performance

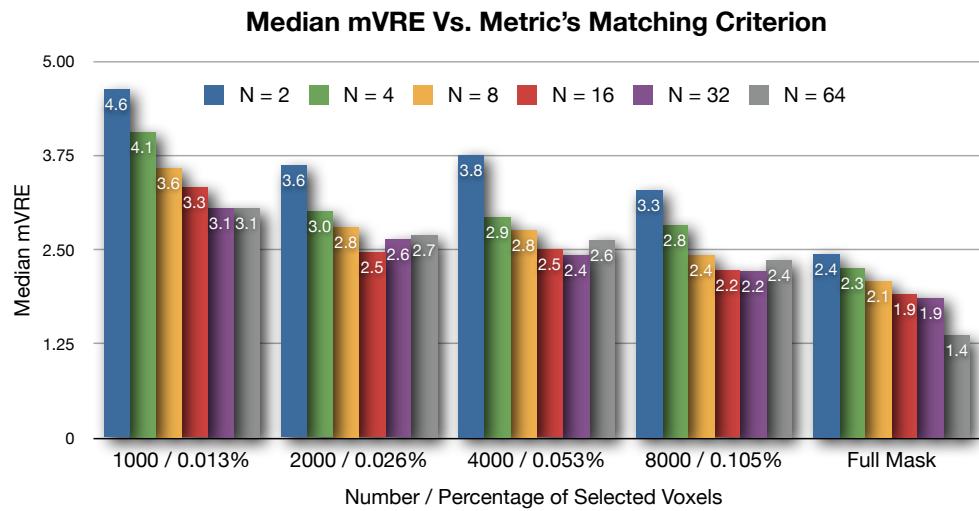


Fig. 5.2 Median mVRE evaluated with varying values of N , see Equation 5.1, and using randomly selected subsets of voxels within a sampling mask defined by the top 10% voxels with highest gradient magnitude. Notice that increasing the selection criterion, tends to yield improvements in registration performance. Also notice, that much of the registration can be resolved with just a few thousand voxels.

gain provided by incorporating additional voxels is relatively small. Finally, it is important to note that the performance obtained with 8,000 voxels (0.105% of the image domain) of the top 10% mask is already very close to the performance obtained with the full image domain and takes on average 28 seconds to compute (in comparison to 1045 seconds when using the full image domain).

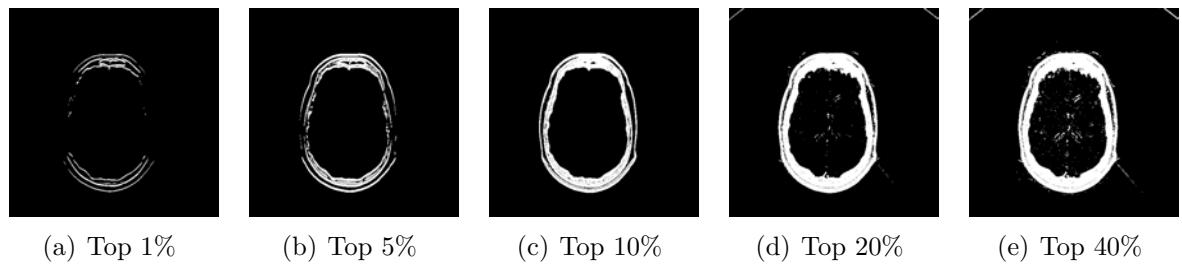


Fig. 5.3 Sampling masks obtained from top percentage of high gradient magnitude voxels.

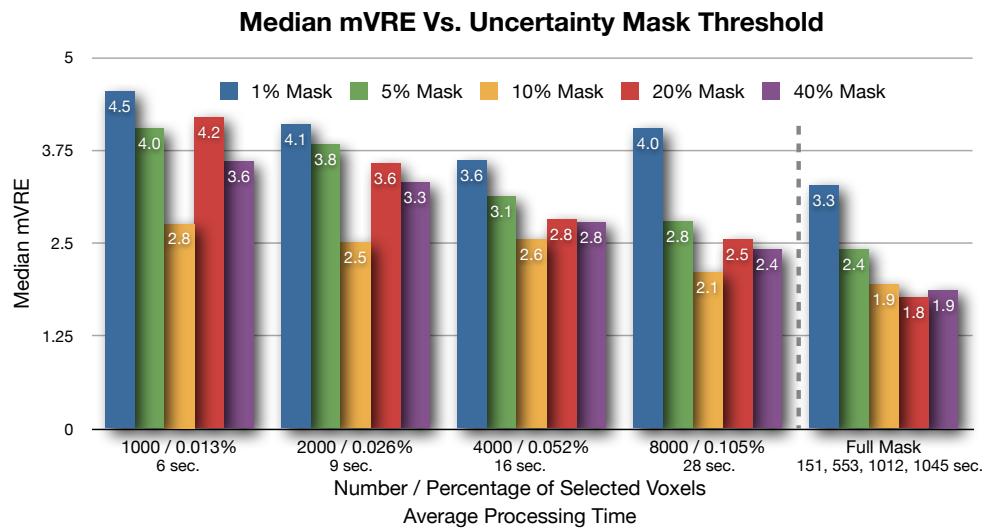


Fig. 5.4 Median mVRE evaluated with different sampling masks obtained from varying thresholds and using randomly selected subsets of voxel within each mask. Notice that when using a subset of voxels, the mask defined by the top 10% voxels consistently yields the best results.

5.4.4 Performance Comparison with Alternative Approaches

The performance of the proposed method is then evaluated across all CT to MRI registration cases found in the testing dataset. The evaluated configuration involves an N value of 32 and a sampling mask defined by the top 10% of high gradient magnitude pixels. In order to demonstrate the trade-off between processing time and registration accuracy, the performance with a reduced number of selected voxels is also evaluated.

Nbr. of Voxels from 10% Mask			1000	2000	4000	8000	All
Fixed	Moving	Nbr. of Cases					
CT	T1	16	1.19	2.08	1.92	2.13	1.63
CT	T2	16	1.59	1.79	1.65	1.69	1.56
CT	PD	11	2.27	2.37	1.85	1.49	1.61
CT	T1-rect	6	1.20	0.98	0.83	0.70	0.64
CT	T2-rect	7	1.26	1.14	0.75	0.87	0.76
CT	PD-rect	7	1.03	0.67	0.78	0.56	0.51

Table 5.1 Rigid registration results expressed in median registration error (in mm) of the VOIs of all cases found in RIRE dataset.

Table 5.1 lists the results of the proposed method. Notice that the performance metric of each task is computed based on the Target Registration Error (TRE) evaluated at each center location of six to ten Volumes of Interests (VOI) per case. The VOIs were identified in the original RIRE project to be of diagnostic and/or surgical importance. Since the VOIs will generally be closer to the center of rotation of the rigid transformation, a performance metric based on VOIs will tend to have smaller values compared to the performance metric used for the training case. The performance for each particular moving image modality is computed as the median mTRE across all cases. Note that the tasks involving rectified MRI volumes show a significantly improved accuracy compared with their non-rectified counterparts. The proposed method shows sub-voxel registration accuracy across all modalities and achieves sub-millimeter accuracy for rectified MRI volumes.

For comparison purposes, in Table 5.2 the performance of alternative top performing methods evaluated with the RIRE dataset is also shown. The most recently published top performing method (NMI_{CUR}) involves the maximization of NMI evaluated over prudently selected voxels which provide highly informative structural information[57] identified with

the use of curvelets. The table also includes the top performing method (NMI_{CP}) previously presented by Collignon et al.[35] and found in the original RIRE publication[172], which also corresponds to the maximization of NMI. Additionally, the results for an alternative technique based on NMI maximization with Generalized Partial Volume Interpolation NMI_{GPV} [30] are also shown. It is important to note that the results from competing methods have been provided by the original authors of each technique and can be found either in the RIRE website or the original publications. Since most methods present only results for cases involving rectified MRI volumes, the comparison of performance will be limited to this subset of cases. Also note that for competing methods evaluated with varying parameterizations, only the results corresponding to their best performing setup are listed. The results shown for the reference top performing method (NMI_{CUR}) correspond to the use of a sampling mask that extends to 15% of the full image domain. Alternatively, the results shown with the proposed method correspond to a sampling mask covering 10% of the image. Notice that the results obtained with the proposed approach have the smallest median values but are similar to the results obtained with NMI_{CUR} . To determine if there is a significant difference between the two techniques, a general linear model repeated measures analysis is performed, where the two techniques (NMI_{CUR} and proposed approach) correspond to within-subject factors and the case is the between-subject factor. The statistical analysis yielded no significant differences between both methods ($F(1,19)=3.502$, $p=0.063$). However, even if there is no significant difference in registration accuracy, the proposed approach used only 10% of the voxels, as opposed to 15% of the voxels, to achieve such performance.

	NMI_{CP}	NMI_{GPV}	NMI_{CUR}	Proposed
	Intensity			Orientation
CT/T1-rect	0.7	0.87	0.66	0.64
CT/T2-rect	0.8	1.11	0.77	0.76
CT/PD-rect	0.8	0.90	0.69	0.51

Table 5.2 Rigid Registration Results expressed in median registration error (in mm) of the VOIs of all cases found in RIRE dataset.

A major motivation of the proposed technique is to exploit the sparse selection of evaluated locations so as to significantly reduce computational times. The following experiments evaluate the registration performance of the proposed framework with extremely aggressive

sampling rates, and compare the results with a previously proposed state-of-the-start pixel selection method. Fig. 5.5 shows the mTRE box plots obtained from all cases with rectified MRI volumes and for different sampling rates. The figure also includes the performance of the top performing (NMI_{CUR}) method with the use of all voxels within a sampling mask defined by either 5% or 10% of the voxels found in the image. The figure also reports the performance of a related NMI method that randomly selects a subset of voxels (i.e. voxels are sampled following a uniform distribution). Notice that the proposed approach is able to resolve the registration with sub-voxel accuracy with extremely few voxels. In particular, one can observe that even when only 0.05% of the voxels are used, one can already obtain a significantly improved registration performance ($F(1,19)=15.755$, $p<0.001$) compared to the one obtained with the top performing NMI method with 5% of the voxels. Similarly, one can observe a significantly improved performance comparing both methods when using 10% of the voxels ($F(1,19)=130.612$, $p<0.001$). The use of a reduced subset of the image brings forward a significant reduction in computation times, since the computational complexity per optimization iteration relates linearly to the number of voxels evaluated. One should note that the stopping criterion for the optimization is not a fixed number of optimization iterations, but rather an arbitrarily small difference in the metric's value or the candidate transformation parameters. Thus, the total computational complexity does not hold a strict linear relation with regards to number of voxels used, but does follow a monotonic relationship.

The idea of relying on image gradient information for multi-modal registration context was previously developed in the work by Pluim et al.[119]. In such work, the authors adopt a similarity measure defined as the product of a gradient-based metric and the mutual information between image intensities. Specifically, the proposed similarity metric from such work is expressed as,

$$S_g(I_f, I_m^\downarrow) = G(I_f, I_m^\downarrow) \cdot S_g^{MI}(I_f, I_m^\downarrow) \quad (5.2)$$

where S_g^{MI} is a MI-based similarity metric (either MI or NMI), and $G(I_f, I_m^\downarrow)$ is a gradient based similarity measure defined as

$$G(I_f, I_m^\downarrow) = \sum_{\mathbf{x} \in \Omega} \left(\frac{\cos(2\Delta\theta) + 1}{2} \right) \cdot \min(|\nabla I_f(\mathbf{x})|, |\nabla I_m^\downarrow(\mathbf{x})|) \quad (5.3)$$

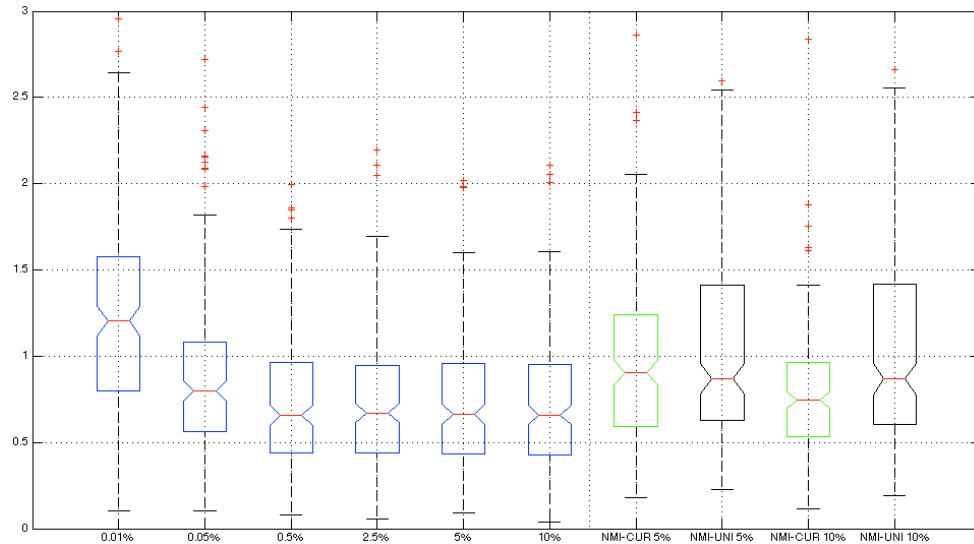


Fig. 5.5 Box plots showing the VOI registration error of 20 different cases of CT to rectified MRI rigid registration with varying number of voxels. The blue box plots to the left of the dotted line correspond to results obtained with the proposed approach and with a varying percentage of evaluated voxels. The green box plots corresponds to the reference top performing method, NMI_{CUR} , consisting in the maximization of NMI evaluated over prudently selected voxels[57]. The black box plots correspond to a method that maximizes NMI evaluated over randomly selected voxels. Notice that with just 0.05% of the voxels one can already obtain more accurate results than those obtained with top performing method with 5% of the voxels and comparable to those obtained with 10% of the voxels with statistical significance ($F(1,19)=15.755$, $p<0.001$)

where $\Delta\theta = \angle(\nabla I_f, \nabla I_m^\downarrow)$ is the inner angle between ∇I_f and ∇I_m^\downarrow .

In other words, the metric is a conventional MI metric weighted by a measure of both gradient orientation alignment and the inferior gradient magnitude from the two images. There are some important differences between the proposed framework found in this thesis and the previously proposed technique. In particular, the previously proposed technique involves the evaluation of a similarity metric over the entire image domain, while the proposed framework relies on the sparse selection of locations which lead to important reductions in processing times. Additionally, the previously proposed technique employs image gradients *in conjunction* with MI, while the proposed framework relies exclusively on image gradient orientation alignment, and therefore involves a further reduced computational cost for

evaluating the metric.

It is thus of particular interest to compare the differences in performance of the proposed approach with such previously presented work. In the original publication [119], the authors measure the accuracy of their proposed metric by setting the initial position to the gold standard solution and then evaluating the position after optimization of the metric. Thus, as stated by the authors, such measures of accuracy are not directly comparable to the registration performance reported by other methods. Nonetheless, they provide an indication as to what might be the best accuracy achievable with methods based on such metrics. Table 5.3 shows the results obtained with such a setup, as found in the original publication, and with four different metrics: MI, NMI, and the product of either MI or NMI and their proposed gradient-based measure. As noted by the original authors, the four evaluated metrics have a similar accuracy with no significant statistical difference. Notice that the proposed approach yields a similar accuracy when actually solving the registration problem from its original position. Furthermore, it manages to achieve such accuracy with a reduced number of voxels as opposed to the full image domain. In fact, the proposed method relies *solely* on the gradient orientation alignment of the top 10% of the voxels with highest gradient magnitude.

	MI	NMI	$MI \times G$	$NMI \times G$	Proposed
CT/T1-rect	0.71	0.61	0.78	0.8	0.64
CT/T2-rect	0.72	0.63	0.89	1.04	0.76
CT/PD-rect	0.68	0.71	0.69	0.87	0.51

Table 5.3 Rigid registration results expressed in median error (in mm) of the VOIs of all cases found in the RIRE Dataset. Results involving intensity based metrics (MI, NMI) and a gradient-based metric (G) are the ones reported by Pluim et al.[119], which are not directly comparable to other methods since their initial position was set to the gold standard solution. The metrics labelled $MI \times G$ and $NMI \times G$ involve the product of the gradient-based metric (G) with either MI or NMI. Note that the gradient-based metric (G) evaluates both gradient magnitude similarity and gradient orientation alignment and is evaluated over the full image domain, whereas the proposed approach is solely based on gradient orientation alignment of the top 10% locations with highest gradient magnitude.

Finally, in Table 5.4, the RIRE identification numbers of the results obtained with the proposed approach are provided for reference. Note that such results involve the CT to

Method ID	Percentage / Number of Voxels
10556	10% / between 7×10^5 and 8.9×10^5
10607	$\sim 5\%$ / 320000
10606	$\sim 2.5\%$ / 160000
10579	$\sim 0.5\%$ / 32000
10578	$\sim 0.1\%$ / 8000
10577	$\sim 0.5\%$ / 4000
10576	$\sim 0.03\%$ / 2000
10575	$\sim 0.01\%$ / 1000

Table 5.4 Identification Numbers for RIRE website corresponding to evaluations of the proposed approach configured with $N = 32$.

rectified MRI dataset, and were characterized by an N value of 32 and a varying number of selected locations.

5.5 Conclusions

In this chapter, I have presented a comprehensive quantitative evaluation of rigid registration performance in a rigid registration context involving CT and MRI brain volumes. The experiments were performed with the use of a publicly available and commonly used dataset, thereby allowing a non-biased direct comparison with previous work.

The first set of experiments were performed so as to characterize the sensitivity of the proposed approach to different parameterizations. I demonstrated that even though the proposed technique yields a robust performance with varying parameter values, there are various components that tend to provide an improvement in accuracy. The two factors that were shown to bring forward such improvements are the use of a sub-voxel location sampling strategy and the use of a highly selective metric where only tightly aligned gradient orientations are rewarded.

The second set of experiments illustrate the performance of the proposed approach with regards to previous work. There are two critically important results. The first result is that the proposed technique can yield a registration accuracy which has no statistical difference compared to the top performing alternative technique, which relies on NMI as a driving similarity metric. The second result is that the proposed approach allows for much more aggressive sampling rates that continue to yield a sub-voxel registration accuracy. In other

words, the registration accuracy of the proposed technique degrades with relation to the number of locations used at a much slower rate than previously reported voxel sampling strategies for MI-based registration.

Chapter 6

Automatic Segmentation based on the Registration of Inter-Patient MRI Brain Volumes

The two previous chapters focused on challenging multi-modal registration contexts that rely on a rigid spatial mapping between images. Such contexts involved a compactly characterized and highly constrained transformation function with a small number of degrees of freedom. The results obtained with the proposed technique in such contexts demonstrated a robust and accurate registration performance while employing extremely reduced sampling rates. Consequently, the proposed technique was shown to provide an important solution for time-sensitive contexts (e.g. intra-operative applications) where reduced processing times can lead to a more seamless and successful clinical integration.

There are many medical contexts that compel the use of a non-rigid deformation model with significantly increased degrees of freedom. Such contexts are commonly addressed with a gradient based optimization strategy (as opposed to an exhaustive non-gradient based technique) so as to maintain the computational complexity within a practical range. Furthermore, the large number of degrees of freedom may also lead to an increased sensitivity to aggressive sampling rates. As such, the type of performance gains shown in the previous chapters may not be obtained in such applications.

This chapter focuses on the application of brain mapping, where a diverse set of anatomical structures are inferred from a patient's medical image. In this context, the preferred

approach for obtaining the anatomical labels is to rely on manual annotations from trained neuroanatomists. However, manual labelling is a slow and costly process that is simply impractical for time-sensitive applications or high-throughput scenarios. The widespread alternative solution adopted by the research community [10, 37, 70, 141, 174] has been to rely on non-rigid image registration techniques which identify the spatial mapping between a patient’s image and a reference image with known anatomical labels. The patient’s anatomical labels can then be directly obtained by finding the spatially corresponding label in the reference image.

This chapter presents the experimental results regarding the proposed diffeomorphic registration strategy, described in Section 3.2, in the context of inter-patient MRI brain registration. This constitutes a mono-modal scenario, as opposed to the multi-modal scenarios seen in the previous chapters. However, the increased number of degrees of freedom can quickly lead to substantial processing times, and there is a strong interest in minimizing computational complexity, while maintaining an accurate registration performance.

The chapter is structured as follows. Section 6.1 describes the publicly available dataset used for the evaluation of inter-patient MRI brain volumes, which involves 12 subjects with the manual labelling of 128 unique brain regions. Section 6.2 then presents a variety of quantitative results obtained with the proposed technique. In particular, Section 6.2.1 first presents a set of diverse results, which characterize the performance of the proposed technique with different parameterizations, and provides a quantitative analysis of how each parameter affects the trade-off between registration accuracy and computational complexity. Section 6.2.2 then presents a direct comparison with 14 state-of-the-art algorithms based on the statistical overlap of labelled brain regions obtained from the different registration techniques. The quantitative comparison shows that the proposed registration technique achieves a registration accuracy similar to the top state-of-the-art algorithms with highly reduced processing times ranging between two and six minutes, and where competing algorithms involve processing times of approximately half an hour. Section 6.2.2 also presents a detailed analysis of segmentation overlap accuracy for each of the 128 unique brain regions. Section 6.2.3 then evaluates the performance of the proposed technique with an alternative edge detection technique so as to illustrate the sensitivity with relation to inferred boundary locations. Section 6.2.4 illustrates the performance of the proposed technique with a varying number of selected voxel locations. Finally, Section 6.3 outlines a series of conclusions synthesizing the results shown in this chapter.

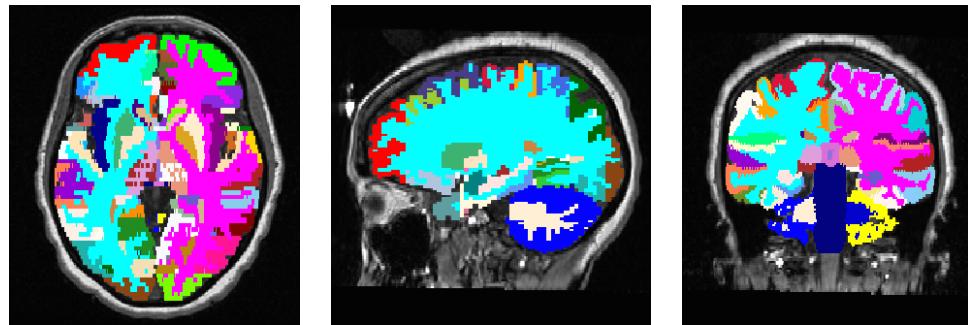


Fig. 6.1 Expert segmented brain regions from the first case of the CUMC12 dataset [84].

6.1 Clinical Data

The proposed approach is evaluated in the context of the non-linear registration of inter-patient MRI brain volumes. In particular, registration performance is evaluated with a public dataset, referred to as the CUMC12 dataset [84], composed of 12 subjects with corresponding manual segmentations of 128 unique brain regions. Fig. 6.1 illustrates the manually segmented brain regions found in the first case of the dataset. The dataset has been used in previous publications [84, 122] for the evaluation of multiple state-of-the-art non-linear registration techniques and thus allows for direct comparison with previously reported results. The volumes found in the dataset were acquired at the Columbia University Medical Center on a 1.5 T GE scanner. For the purposes of manually segmenting the volumes, the original images were resliced coronally to a slice thickness of 3 mm. The resolution of the manual labels is thus slightly coarser than the resolution of the native MRI volumes, which have a coronal slice thickness of 1.5 mm. The expert labellers followed the Cardviews labeling protocol [26] with the use of the Cardviews software.

Registration performance is evaluated by applying the estimated non-linear transformation to the expert labels of the moving image, and then comparing the warped labels with the expert labels of the fixed image. The statistical agreement between labels is then evaluated with the target overlap measure and the union overlap measure [84].

6.2 Experimental Results

This section presents a series of experimental results which illustrate the sensitivity of the proposed technique to different parameterizations of the transformation model, compare the performance of the proposed technique to multiple competing state-of-the-art methods, evaluate the performance of the technique with a varying number of selected locations, and demonstrate the performance of the technique when relying on an alternative edge detection technique. Note that the approach was implemented in C++ and evaluated on a Linux computer with an Intel i7-3770 CPU.

6.2.1 Sensitivity to Spatial and Temporal Discretization

The diffeomorphic registration framework outlined in Section 3.2 involves various transformation model parameters, which ultimately define the degrees of freedom of the diffeomorphic transformation. This section presents a quantitative evaluation of registration performance with relation to varying parameterizations of the transformation model. The goal is to provide an assessment of the trade-off between computational complexity and registration accuracy. It is important to note that the results shown in this section involve only one third of the cases found in the CUMC12 dataset, given that the experiments are focused on relative variations in performance of the proposed technique. In particular, the first four brain volumes of the dataset are registered against all other brain volumes.

The first performance evaluation involves the use of a static (i.e. time-constant) velocity field and compares the mean segmentation overlap with the average processing time across registration cases. The transformation model is parameterized with varying voxel spacings for the velocity field image and with varying number of time steps (i.e. time-discretization of the diffeomorphism). Note that coarser voxel spacings for the velocity field image lead to a reduction in the degrees of freedom of the transformation, and essentially restrict the transformation variation within a voxel's volume. Also note, that a small number of time steps not only limits the flexibility of the transformation model, but may also lead to significant inconsistencies between the forward and the backward transformation.

Fig. 6.2 illustrates the results obtained from such experimental setup with two different visualizations represented as bubble charts. In the top figure, the bubble's color reflects a unique value of the voxel spacing, and the bubble's size and value reflect the number of time steps involved. In the bottom figure, the bubble's color reflects a unique value of the

number of time steps involved, and the bubble's size and value reflect the voxel spacing. Both visualizations allow for a simple analysis of the trade-off between registration accuracy and processing time for different parameterizations, where bubbles farther up and left reflect an improved trade-off. The results found in Fig. 6.2 show that the increase in temporal resolution obtained by increasing the number of time steps leads to a comparable increase in computational complexity, but seems to provide little gain in terms of segmentation overlap. Alternatively, the increase in spatial resolution (i.e. decreased voxel spacing) of the velocity field leads to significant gains in segmentation overlap while incurring in slightly increased processing times.

The second performance evaluation follows the same pattern but adopts a dynamic (i.e. time-dependent) velocity field. In particular, it also involves the comparison of the mean segmentation overlap with the average processing time across registration cases with varying voxel spacings and number of time steps.

Fig. 6.3 illustrates the results obtained from such experimental setup with two different visualizations represented as bubble charts. Similar to the results found in Fig. 6.2, the results in Fig. 6.3 show that the increase in temporal resolution, obtained by increasing the number of time steps, leads to a comparable increase in computational complexity, but seems to provide little gain in terms of segmentation overlap. Furthermore, an increase in spatial resolution (i.e. decreased voxel spacing) of the velocity field leads to significant gains in segmentation overlap while incurring in slightly increased processing times. It is important to highlight that even the coarsest discretization of the transformation model already provides a reasonable mean target overlap of 48.4% with an average processing time of 75 seconds, whereas the finest discretization of the transformation model provides an improved mean target overlap of 52.1% with an average processing time of 454 seconds. In other words, a slight increase in mean target overlap comes with a six-fold increase in computational complexity. The following section illustrates how the results obtained with the proposed technique compare with alternative state-of-the-art methods.

6.2.2 Comparison to other methods

In this section, the performance of the proposed approach is compared against alternative state-of-the-art methods. To allow a direct comparison with previously reported results, the proposed technique is evaluated across all cases found in the CUMC12 dataset. The

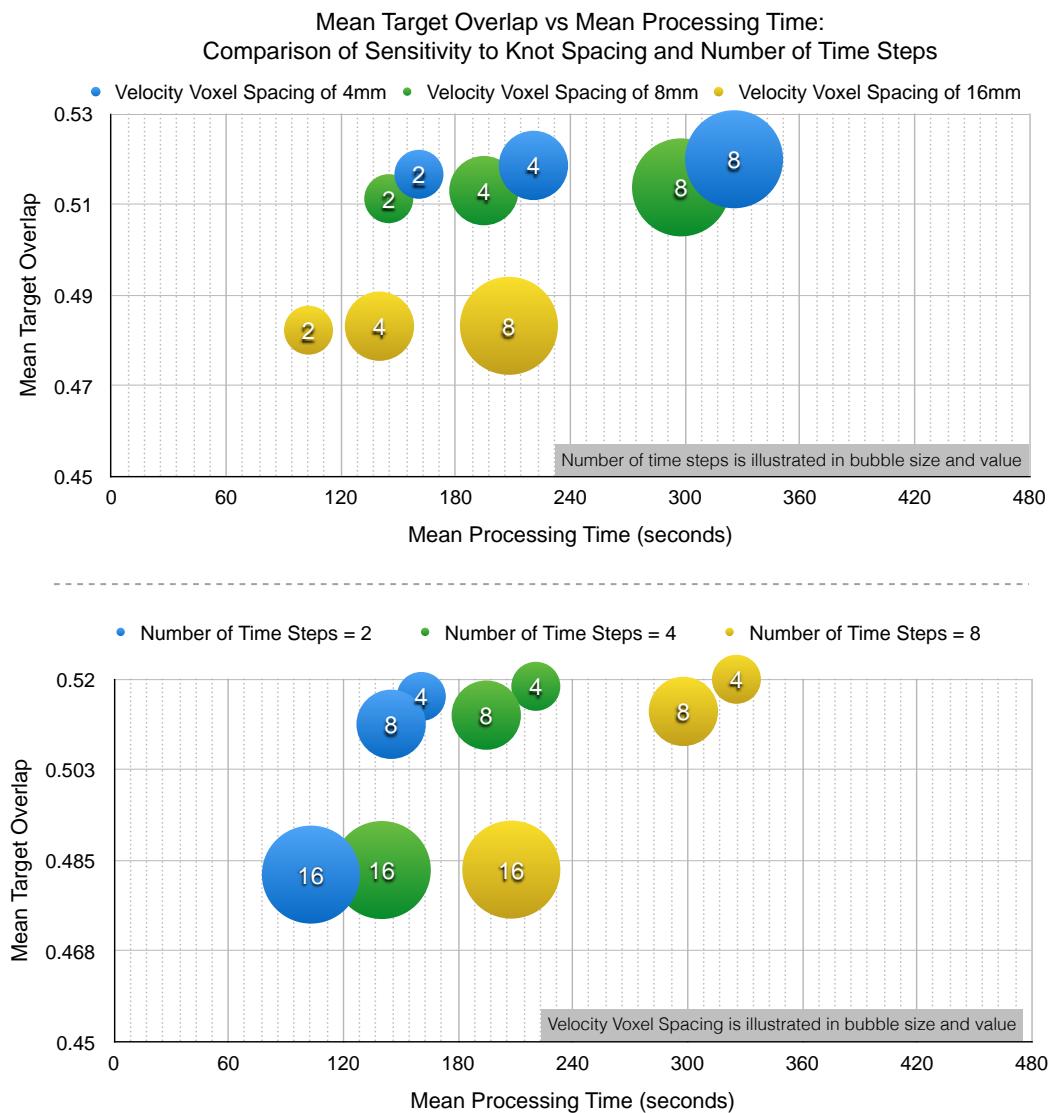


Fig. 6.2 Comparison of segmentation overlap and processing time for a static velocity field with varying values of velocity voxel spacing and number of time steps. The top and bottom figure evaluate the segmentation overlap with varying values of the number of time steps used to temporally discretize the diffeomorphism and with varying values of the voxel spacing of the velocity field image. The y-axis corresponds to the mean target overlap across all regions and across one third of the registration cases from the CUMC12 dataset. The x-axis corresponds to the mean processing time across cases.

proposed approach is parameterized with four coarse-to-fine registration stages, where at each consequent stage the image resolution as well as the resolution of the velocity field is

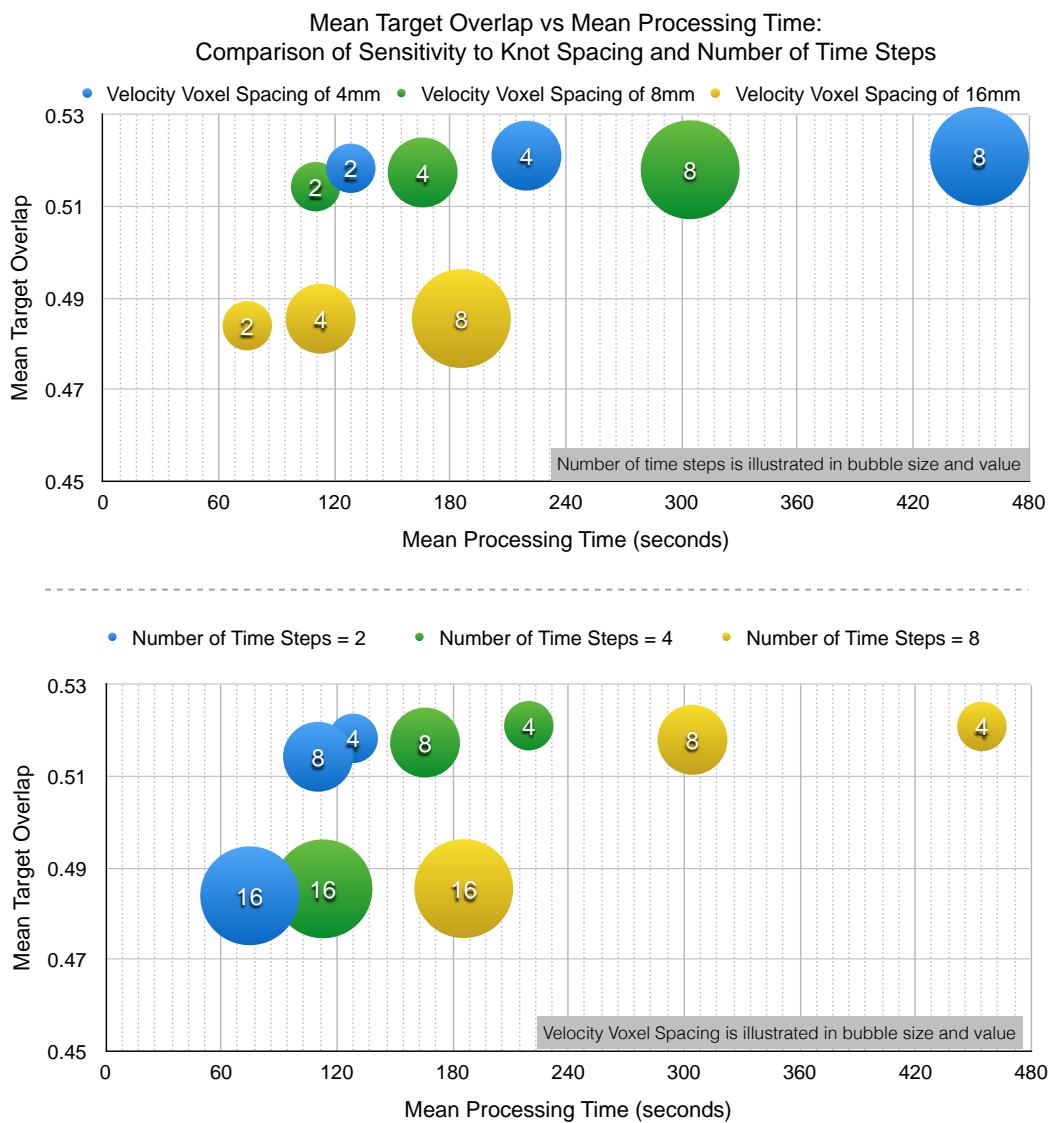


Fig. 6.3 Comparison of segmentation overlap and processing time for a *dynamic* velocity field with varying values of velocity voxel spacing and number of time steps. The top and bottom figure evaluate the segmentation overlap with varying values of the number of time steps used to temporally discretize the diffeomorphism and with varying values of the voxel spacing of the velocity field image. The y-axis corresponds to the mean target overlap across all regions and across one third of the registration cases from the CUMC12 dataset. The x-axis corresponds to the mean processing time across cases.

increased. The finest registration stage involves a velocity field with a voxel spacing of $4 \times 4 \times 4$ mm, and all registration stages are parameterized with four time steps. Registration

performance is evaluated across all stages and is summarized in Tables 6.1 and 6.2. The results presented on Table 6.1 rely on a static (i.e. time-constant) velocity field, while the results presented on Table 6.2 rely on a dynamic (i.e. time-dependent) velocity field. Note that the processing time for each stage includes the processing time of the preceding stages. Hence, the complete registration with all four stages takes 4 minutes and 30 seconds and 6 minutes and 30 seconds respectively. Also note that the gain in registration accuracy between stage three and four is relatively small, and that a registration with only the first three stages takes on average 1 minute and 12 seconds or 1 minute and 30 seconds, depending on the type of velocity field used.

Stage	Mean Target (%)	Mean Union (%)	Cumulative Run Time
1	41.26	26.30	7 secs
2	46.94	30.94	21 secs
3	50.32	33.88	1 min 12 secs
4	51.40	34.65	4 min 30 secs

Table 6.1 Registration performance and processing time for each registration stage with a static velocity field. Performance is evaluated with the mean target and mean union (Jaccard) overlap measure across cases and regions. Note that the processing time for each stage includes the processing time of the preceding stages.

Stage	Mean Target (%)	Mean Union (%)	Cumulative Run Time
1	41.19	26.24	6 secs
2	47.17	31.12	26 secs
3	50.69	34.18	1 min 30 secs
4	51.65	34.75	6 min and 30 secs

Table 6.2 Registration performance and processing time for each registration stage with a dynamic velocity field. Performance is evaluated with the mean target and mean union (Jaccard) overlap measure across cases and regions. Note that the processing time for each stage includes the processing time of the preceding stages.

The results of the proposed approach, referred to as *SymBA*, can be directly compared with the results reported on previous work [84], which focused on the exhaustive comparison of 14 non-linear registration techniques. The proposed technique is configured with either



Fig. 6.4 Compound mean target and union (Jaccard) overlap across regions and across cases. The proposed approach (in red), *SymBA*, is within the top five techniques, and is configured with either a static velocity field, referred to as *SymBA_S*, or with a dynamic velocity field, referred to as *SymBA_D*.

a static velocity field, referred to as *SymBA_S*, or with a dynamic velocity field, referred to as *SymBA_D*. Fig. 6.4 shows the compound average of the target overlap and union overlap (Jaccard) across regions and across registration cases. Note that the proposed approach has similar results to the top five techniques, and all such methods have comparable performance. The union overlap results found in Fig. 6.4 also include the numbers reported on a recently proposed registration technique [122], referred to as *FEM*, that involves an average processing time of 15 minutes, as reported by the authors, which is approximately 3 times the average processing time of the proposed technique.

Segmentation Overlap per Anatomical Region

The compound means for each overlap measure are useful global metrics for comparing performance across methods, yet they do not describe the variability in performance found across regions. Fig. 6.5 illustrates the mean and standard deviation of the target overlap measure for each region. To minimize visual clutter results from only three techniques are

included: a top ranking method from the top five techniques, *SyN* [10]; the proposed approach, *SympBA_S* (with a static velocity field); and an alternative method, *ROME0* [70], which relies on a similarity term based on optical-flow and a regularization term based on the quadratic difference of the deformation field computed between neighboring points. Note that the performance of the proposed approach tends to tightly match the performance of the top ranking method, *SyN*, and is consistently superior to *ROME0*. However, the average processing time of *SyN* was of 38 minutes when ran on the same computer where the proposed technique was implemented, as compared to the proposed technique which takes approximately 4 minutes and 30 seconds.

6.2.3 Comparison with Alternative Edge Detection

The proposed technique has an explicit dependency on an edge detection strategy and hence any variability in edge detection performance will lead to a variability in registration performance. The experimental results presented so far have relied on a Canny edge detection technique. In order to illustrate the variability that is induced by different inferred boundary locations, this section presents the registration performance obtained with an alternative edge detection technique evaluated across all cases found in the CUMC12 dataset. In particular, the alternative edge detection strategy first identifies all the zero crossings of the Laplacian of a Gaussian-blurred image, and then discards all locations whose gradient magnitude falls below a specified threshold.

Table 6.3 illustrates the registration performance obtained over the four registration stages with the two edge detection technique and with a static velocity field. Note that the performance obtained with the alternative edge detection tends to be slightly inferior than the one obtained with the Canny edge detector, but generally achieves a very similar accuracy. The difference in performance may be explained by the fact that the Canny edge detector can identify more subtle boundaries based on the hysteresis thresholding mechanism it employs. Nonetheless, it is important to highlight that the variability in registration performance appears to be relatively minor, which suggests a significant robustness to varying performance in edge detection.

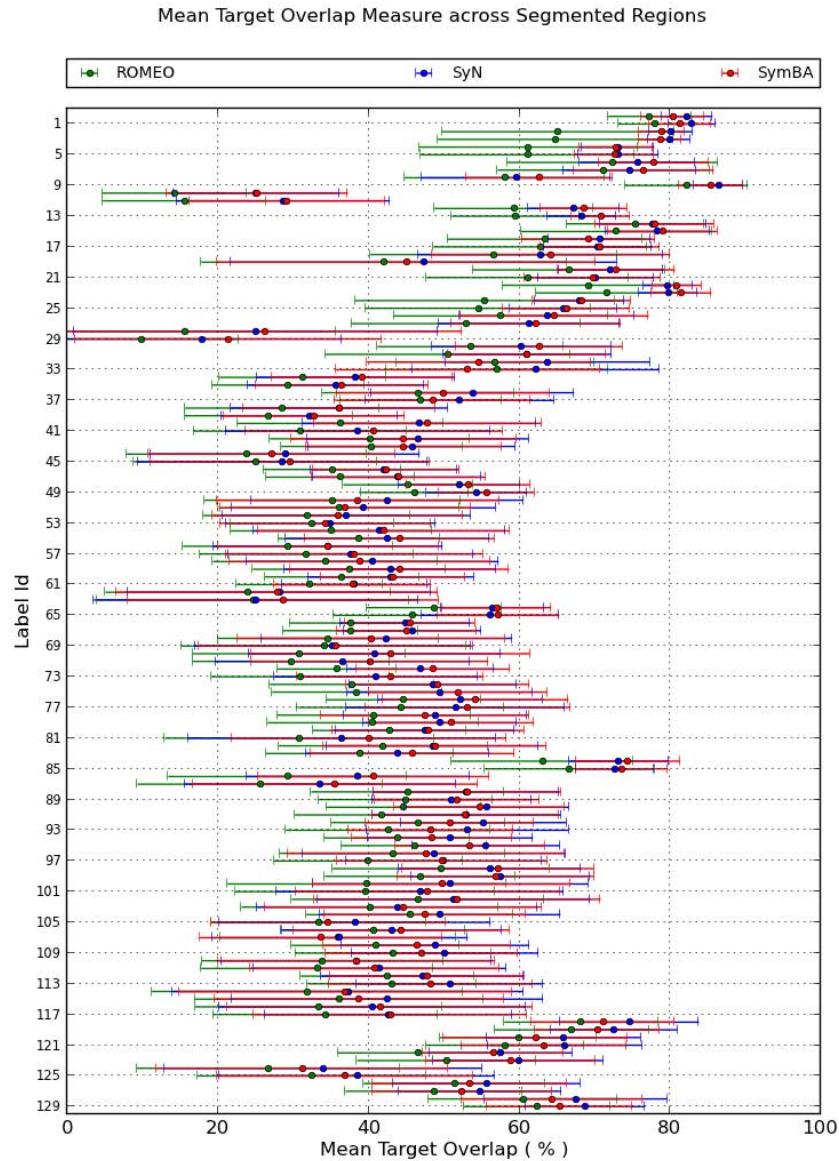


Fig. 6.5 Mean target overlap and standard deviation per region. Registration performance of the proposed method, *SymBA* (with a static velocity field), is compared against three techniques: a top performing method, *SyN*, and *ROMEO*. Note how the registration performance of the top performing method, *SyN*, closely resembles the performance of the proposed method, and is consistently superior to the low ranking method. *SyN* had an average processing time of 38 minutes, while *SymBA* had an average processing times of 4 minutes and 30 seconds when evaluated on the same machine.

Zero-Crossing of Laplacian

Stage	Mean Target (%)	Mean Union (%)	Cumulative Run Time
1	41.48	26.45	6 secs
2	46.56	30.61	22 secs
3	49.75	33.38	1 min and 11 secs
4	50.78	34.07	4 min and 25 secs

Canny Edge Detection

Stage	Mean Target (%)	Mean Union (%)	Cumulative Run Time
1	41.26	26.30	7 secs
2	46.94	30.94	21 secs
3	50.32	33.88	1 min 12 secs
4	51.40	34.65	4 min 30 secs

Table 6.3 Registration performance and processing time for each registration stage with a static velocity field and two alternative edge detection techniques: one based on the zero-crossings of the Laplacian, and the Canny edge detection technique that was used for all other experiments. Performance is evaluated with the mean target and mean union (Jaccard) overlap measure across cases and regions. Note that the processing time for each stage includes the processing time of the preceding stages.

6.2.4 Sensitivity to Sampling Rate

The proposed technique relies on a sparse selection of voxel locations, which substantially reduces the number of voxels that will be evaluated throughout the optimization and leads to a reduction in computational complexity. This section evaluates the performance of the proposed technique when randomly selecting a further reduced number of voxel locations from the set of boundary locations, which can provide even further reduced processing times at the expense of a degradation in registration accuracy.

Figures 6.6 and 6.7 illustrate the mean target overlap of the proposed technique across all cases in the CUMC12 dataset with either a static or dynamic velocity field and with a varying number of evaluated voxel locations. For reference, the figures also include the mean target overlap of five competing methods. It is important to note that the number of evaluated voxel locations specified corresponds only to the final stage of the proposed technique, and all other previous stages have a common parameterization. Note that the

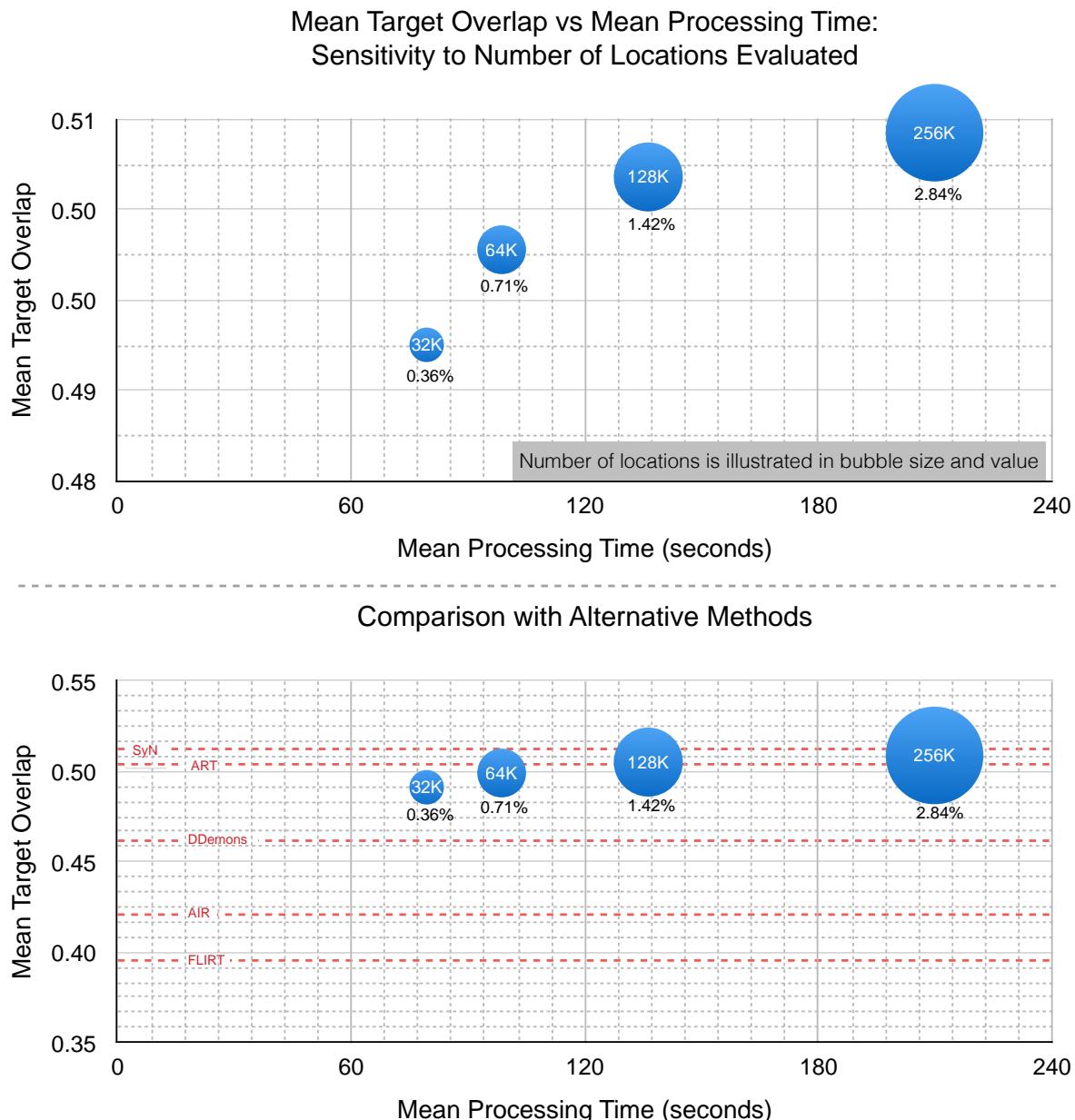


Fig. 6.6 Registration performance and processing time with different sampling rates evaluated with a *static* velocity field. The white labels in the center of each bubble correspond to the number of voxel locations evaluated at each image, and the black labels below each bubble reflect the corresponding percentage of voxels selected (with relation to the total number of voxels in the image). The bottom figure includes the mean target overlap obtained with five competing methods that serve as a reference for state-of-the-art registration accuracy.

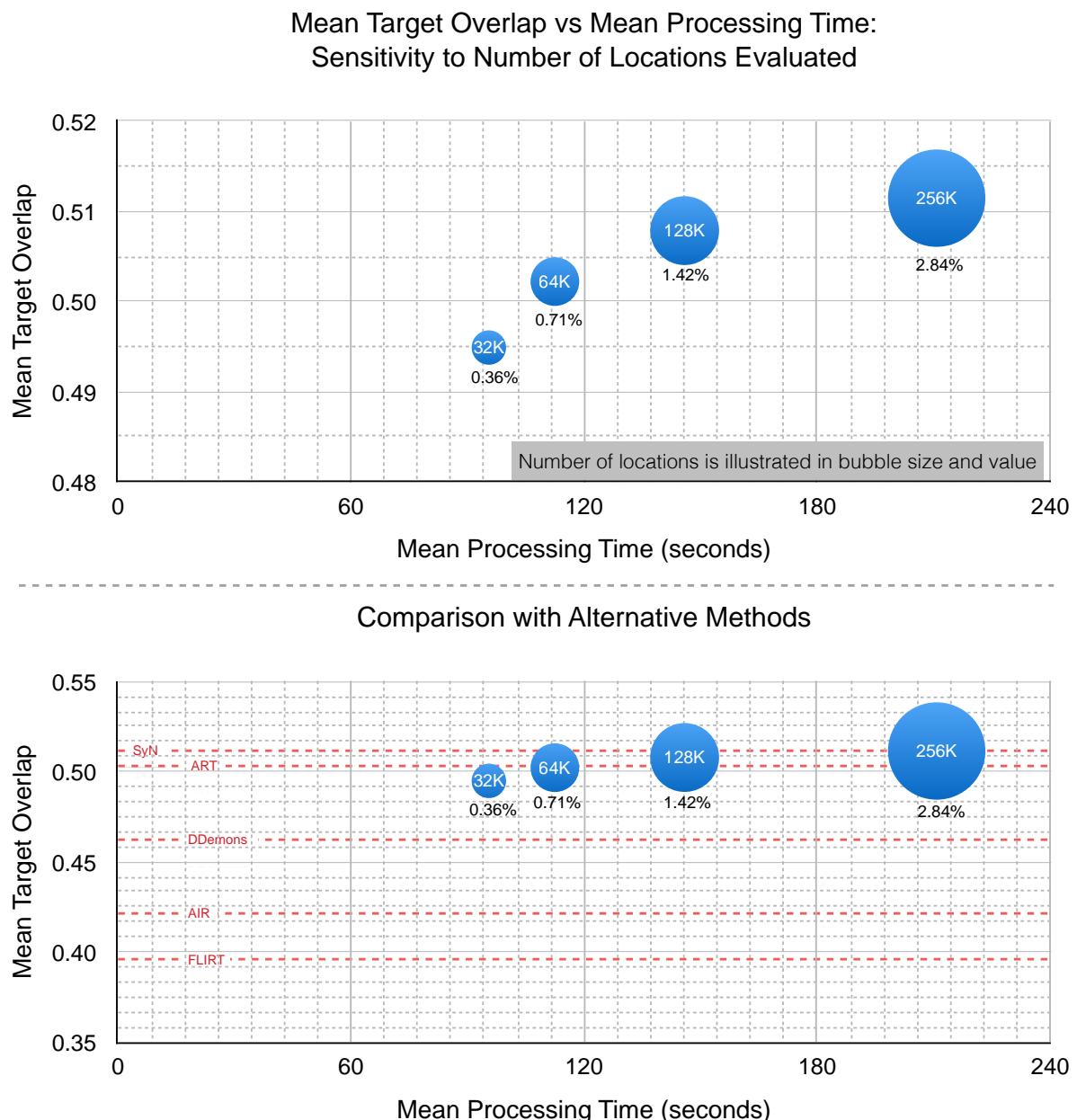


Fig. 6.7 Registration performance and processing time with different sampling rates evaluated with a *dynamic* velocity field. The white labels in the center of each bubble correspond to the number of voxel locations evaluated at each image, and the black labels below each bubble reflect the corresponding percentage of voxels selected (with relation to the total number of voxels in the image). The bottom figure includes the mean target overlap obtained with five competing methods that serve as a reference for state-of-the-art registration accuracy.

gain in registration accuracy obtained in terms of the increase in the number of voxel locations (and thus processing time) exhibits diminishing returns. In other words, the degradation in registration accuracy with reduced sampling rates is relatively small. In particular, with only 64,000 voxel locations (which corresponds to approximately 0.7% of the voxels in the entire volume) and a dynamic velocity field, the proposed technique achieves a mean target overlap of 50.2%, which places it within the top five methods evaluated in [84] (illustrated in Fig. 6.4). Such performance provides a valuable tool for time-sensitive contexts where a small loss of registration accuracy can be justified for highly reduced processing times.

6.3 Conclusions

In this chapter, I have presented a quantitative evaluation and comparison of the proposed diffeomorphic registration technique in the context of inter-patient brain MRI registration. The experiments were performed with the use of a publicly available used dataset which has been previously used for the evaluation of non-linear registration techniques. Hence, it allows for a non-biased direct comparison with related methods.

One set of experiments was performed so as to characterize the trade-off between computational complexity and registration accuracy based on different parameterizations. The results demonstrate that the proposed method can provide extremely reduced processing times with a minor penalty in registration accuracy. In particular, the evaluation in terms of varying sampling rates showed that even with only 0.7% of the voxels in the image domain, the proposed technique can obtain a mean target overlap of 49.3%, which is already within the accuracy of top 5 methods evaluated in [84].

Additionally, the evaluation in terms of varying number of time steps also showed that increasing the number of time steps only provided minor gains in registration accuracy while substantially increasing processing times. Nonetheless, it is important to highlight that while registration accuracy remains relatively similar for a different number of time steps, a small number of time steps involves a coarse discretization of time that can lead to inconsistencies between the forward and backward transformation. The focus of this thesis is largely on the similarity term of the image registration technique and how it can be computed efficiently, and has not devoted an effort to explicitly ensuring that the transformation model yields an accurately consistent diffeomorphism. In order to ensure consistency of the

diffeomorphic transformation and enforce smoothness in the spatial mapping, further work should consider increasing the number of time steps until the diffeomorphic transformation is consistent (up to a predefined accuracy), as well as incorporating a regularization term, such as the one found in the LDDMM [16] framework, which is typically computed as the L^2 -norm of a differential operator on the velocity field.

A second set of experiments compared the performance of the proposed technique against 14 state-of-the-art approaches and demonstrated that the registration accuracy very closely approaches that of the top performing methods. However, such registration accuracy was obtained with a drastic reduction in processing times. In particular, the average processing time of the proposed technique ranged between 1 and 6 minutes (depending on the configuration), while the average processing time of a top performing method was of 38 minutes.

Such substantial reductions in processing times with a minor penalty in registration accuracy can be fundamental for time-sensitive or high-throughout applications. For example, consider the context of multi-atlas label fusion (MALF) where each scan is registered against a *series* of references atlases with expert anatomical labels, and the spatially corresponding anatomically labels are then *fused* in a smart fashion so as to identify the anatomical label of the patient’s scan. MALF approaches are very popular automatic segmentation strategies that have shown improved accuracy in multiple contexts. However, they are characterized by a very high computational cost, which corresponds directly to the series of registrations against each atlas. In such contexts, the proposed technique can provide a highly valuable tool for either substantially reducing the overall processing time or alternatively increasing the number of reference atlases that can be considered in a fixed amount of time.

The inter-patient registration of healthy MRI brain volumes for automatic anatomical labelling constitutes an important and popular application where the proposed technique has shown to provide important gains. A somewhat related context involves the generation of healthy soft tissue probabilities for each voxel of a patient’s MRI volume via the non-rigid registration of a patient’s brain MRI to template brain MRI (which includes predefined healthy tissue probabilities for each voxel). This application can be found as a fundamental preliminary stage for the automatic detection and segmentation of brain pathologies. The next chapter focuses on the specific context of the automatic segmentation of brain lesions brought forward by multiple sclerosis (MS) and illustrates the improvements in overall

segmentation performance and processing times that can be obtained with the use of the proposed registration technique.

Chapter 7

Tissue Priors via the Registration of Patient to Template MRI Brain Volumes in the Presence of Pathology

The previous chapter provided experimental results relating to the inter-patient non-rigid registration of brain volumes so as to map anatomical labels. A main conclusion from the quantitative comparison with alternative state-of-the-art methods was that the proposed technique achieves highly reduced processing times with a registration accuracy that closely approaches that of the top performing techniques. As such, the proposed technique constitutes an important improvement towards time-sensitive or high-throughput applications where processing times represent a critical operational bottleneck.

This chapter focuses on a closely related context involving the non-rigid registration of patient MRI brain volumes to a standard MRI brain atlas so as to obtain healthy tissue probabilities for each voxel in the patient's volume. In particular, the obtained healthy tissue probabilities are considered in the specific context of brain lesion detection and segmentation, where the probability maps serve as inputs for automatic segmentation algorithms. It is important to note that this context involves brains with pathologies whereas the context in the previous chapter was restricted to healthy brains. This constitutes an important challenge to any registration algorithm since the pathological tissue may easily be misinterpreted as a tissue displacement. Such a challenge is one manifestation of the general limitation of relying on the spatial mapping to a *healthy* brain atlas for inferring

soft tissue priors. In reality, there is simply no one-to-one spatial mapping between the anatomy of the healthy brain atlas and the unique anatomy of the pathological brain volume, and the healthy soft tissue probabilities derived from such process can only provide a coarse, yet critically valuable, indication of what the underlying tissue truly is.

The evaluation of registration performance presented in this chapter is fundamentally different from the typical evaluation of performance which tries to characterize geometric accuracy directly or through some surrogate measure. In particular, the experimental setup in this chapter focuses on the bottom-line performance for this context, which is the brain lesion detection and segmentation performance. Such “down-stream” performance evaluation constitutes an important analysis for modern medical image analysis pipelines that are commonly composed of a series of image processing algorithms. Since each of these processing stages can be performed by one of many competing state-of-the-art techniques, it is of critical interest to characterize the variation in bottom-line performance which can be brought forward by alternative choices for each stage.

The chapter is structured as follows. Section 7.1 briefly describes the clinical context of multiple sclerosis (MS) and the particular benefit brought forward by automatic MS lesion detection and segmentation algorithms. Section 7.2 provides an overview of the set of real clinical data used for evaluation of lesion segmentation performance, which involves ground truth obtained from a meticulous manual labelling process. Section 7.3.1 provides a description of the results referring to Gad lesion detection, which involves lesions associated with active inflammations in the brain that are highlighted by a gadolinium MRI contrast agent. Sections 7.3.2 present the results obtained in the context of T2 lesion detection, which involves chronic lesions that are commonly exposed by hyper-intensities in T1-weighted MRI volumes in white matter regions. In each context, the bottom-line performance obtained with the proposed technique is compared against the one obtained with two competing state-of-the-art registration algorithms. Similar to the results obtained in the previous chapter, the proposed technique is shown to achieve a slightly improved bottom-line performance with significantly reduced processing times.

7.1 Multiple Sclerosis Brain Lesion Detection And Segmentation

The automatic detection of MS brain lesions plays an important role in the context of clinical trials for treatments where detected MS brain lesions serve as biomarkers to see if a

treatment has a statistically significant effect on the population [153]. These clinical trials are commonly composed of hundreds of patients scanned at multiple time-points, and thus involve a high data throughput. The benefit of relying on an automatic detection algorithm over expert manual labelling clearly involves reduced costs and reduced turnaround times. As such, any significant improvement in computational efficiency provides an increased ability towards quickly processing such large-scale experiments.

Algorithms for the automatic detection and segmentation of brain lesions and tumours [51, 81, 158] commonly rely on healthy tissue prior probabilities for facilitating the task of detecting pathologies. Such healthy tissue priors typically include tissue types such as grey matter (GM), cerebrospinal fluid (CSF), and white matter (WM). Patient specific priors are commonly obtained by registering a standard brain template, whose tissue probabilities are known, to the patient's T1-weighted MRI volume. The work presented in this section relies on a brain atlas (i.e. template) known as the ICBM 2009a non-linear symmetric atlas [56], which was generated by the McConnell Brain Imaging Centre in McGill University. Fig. 7.1 illustrates the multiple MRI modalities included in such atlas, along with critically valuable anatomical information. In particular, the ICBM 2009a atlas includes MRI volumes from T1-weighted, T2-weighted, PD-weighted modalities, and T2 relaxometry, as well as tissue probability maps for GM, CSF and WM, and anatomical masks for the brain, eyes, face and distinct brain lobes.

The tissue probabilities are particularly valuable for detecting and segmenting pathologies that arise in specific tissue types. For example, brain lesions brought forward by MS have a strong predilection for periventricular WM and a slight predilection for cortical WM. Hence, any segmentation strategy will typically characterize voxels with high WM probability as having a higher probability of corresponding to a lesion than voxels with a low WM probability.

Unfortunately, the healthy tissue priors obtained via registration are not perfectly accurate and even slight misregistrations of the MRI volume can lead to a degradation in segmentation performance. Consider the example shown in Fig. 7.2 illustrating an expert labelled MS lesion overlayed onto two alternative WM tissue priors. The middle column of Fig. 7.2 shows an accurate registration where the high probability WM voxels have a strong overlap with the expert labelled brain lesion. This constitutes a physiologically consistent soft tissue prior that concurs with the principle that MS lesions are generally identified as scars that affect WM tissue. Alternatively, the right column of Fig. 7.2 shows a slightly

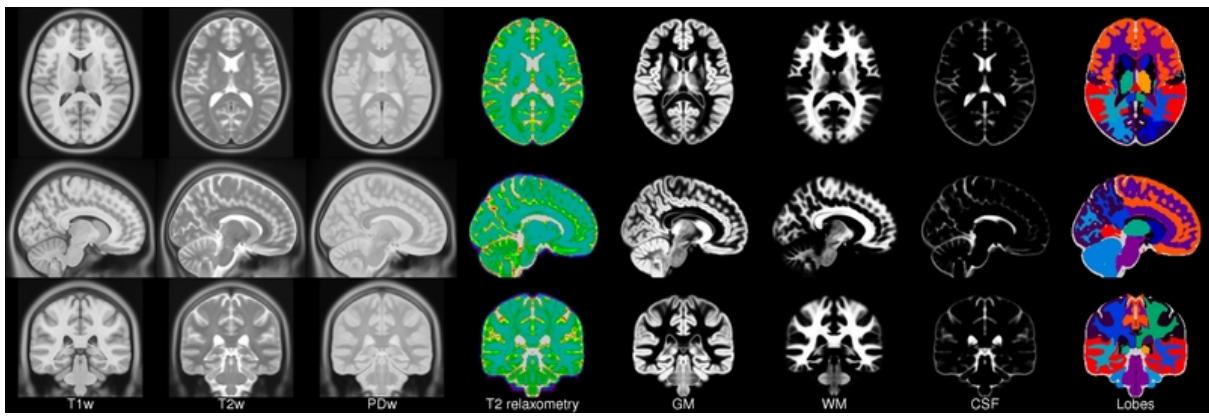


Fig. 7.1 ICBM 2009a brain atlas from McConnell Brain Imaging Centre in McGill University. The atlas includes MRI volumes from T1-weighted, T2-weighted, PD-weighted modalities, and T2 relaxometry, as well as tissue probability maps for GM, CSF and WM, and anatomical masks for the brain, eyes, face and distinct brain lobes. Image obtained from <http://www.bic.mni.mcgill.ca/ServicesAtlases>.

inaccurate registration that leads to a low WM prior in the lesion region, which constitutes a physiologically inconsistent soft tissue prior. Such inconsistency can negatively affect an automatic lesion segmentation algorithm if it is present either in the training dataset or the testing dataset. If it is present in the training dataset it can mistakenly lead to an increased learned probability of MS lesions being found in non WM tissue. Hence, it can consequently trigger an increase in the number of false detections in non WM tissue. Alternatively, if the misregistration is present in the testing dataset, the automatic algorithm will have a reduced likelihood of detecting the lesion.

This chapter evaluates the task of detecting MS lesions, referred to as *T2 lesions*, as well as the task of detecting gad-enhancing MS lesions, referred to as *Gad lesions*, in brain MRI. Furthermore, each task is evaluated with a unique state-of-the-art method [81, 158].

Gad lesions are associated with active inflammations in the brain. They are best visualized in T1-weighted MRI volumes obtained after injection of a gadolinium-based contrast agent, as illustrated in Fig. 7.3. Automatic segmentation of such lesions is challenging since they involve a wide variation of shapes, sizes and locations. Gad lesions also involve a broad voxel intensity range with significant overlap with healthy tissue types, making voxel intensity values insufficient for accurately discerning between lesions and healthy tissue. Such lesions can also be particularly small (as small as three voxels) and thus may involve

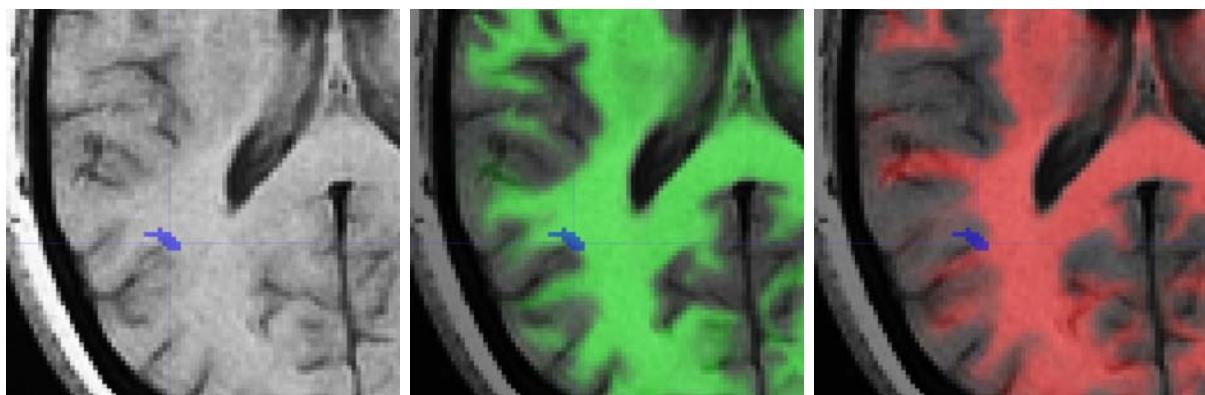


Fig. 7.2 Expert labelled brain lesion overlayed onto alternative WM priors. Left column shows an axial slice of a T1-weighted MRI with an expert labelled brain lesion (in blue). Middle column shows the same slice overlaid with the WM priors obtained via registration with proposed approach (in green). Right column shows the slice overlaid with the WM priors obtained via registration with an alternative approach (in red). Voxels with intense prior coloring (red or green) reflect locations with an increased WM tissue probability. Brain lesions brought forward by MS are generally found in WM tissue. Note how the proposed approach yields an improved overlap of the high WM probability voxels with the lesion, which matches the pathophysiology of the disease. In contrast, the alternative approach yields a weak overlap of high WM probability voxels with the expert labelled lesion, which constitutes a mistakenly strong prior belief that the location cannot be a lesion.

incredibly subtle image differences. Additionally, there are healthy anatomical structures, such blood vessels, which are also enhanced by the contrast agent and may be mistaken for lesions. The spatial distribution of Gad lesions is typically concentrated on either deep WM tissue (e.g. close to lateral ventricles) or near the cortex. Both regions constitute sensitive areas for registration techniques. In particular, periventricular regions of pathological brains may involve enlarged ventricles which involve complex local displacements with respect to the brain atlas and can be challenging to register. Similarly, WM-GM tissue boundaries found in the cortical brain regions are challenging for registration due to the anatomical variability between individual patients and the brain atlas, as well as the subtle intensity difference that characterizes the precise location of the healthy tissue boundary. In short, the automatic detection of Gad lesions is challenged by a variety of factors and thus compels the use of additional information such as the healthy tissue probabilities obtained from registration of the patient's brain MRI to a brain atlas.

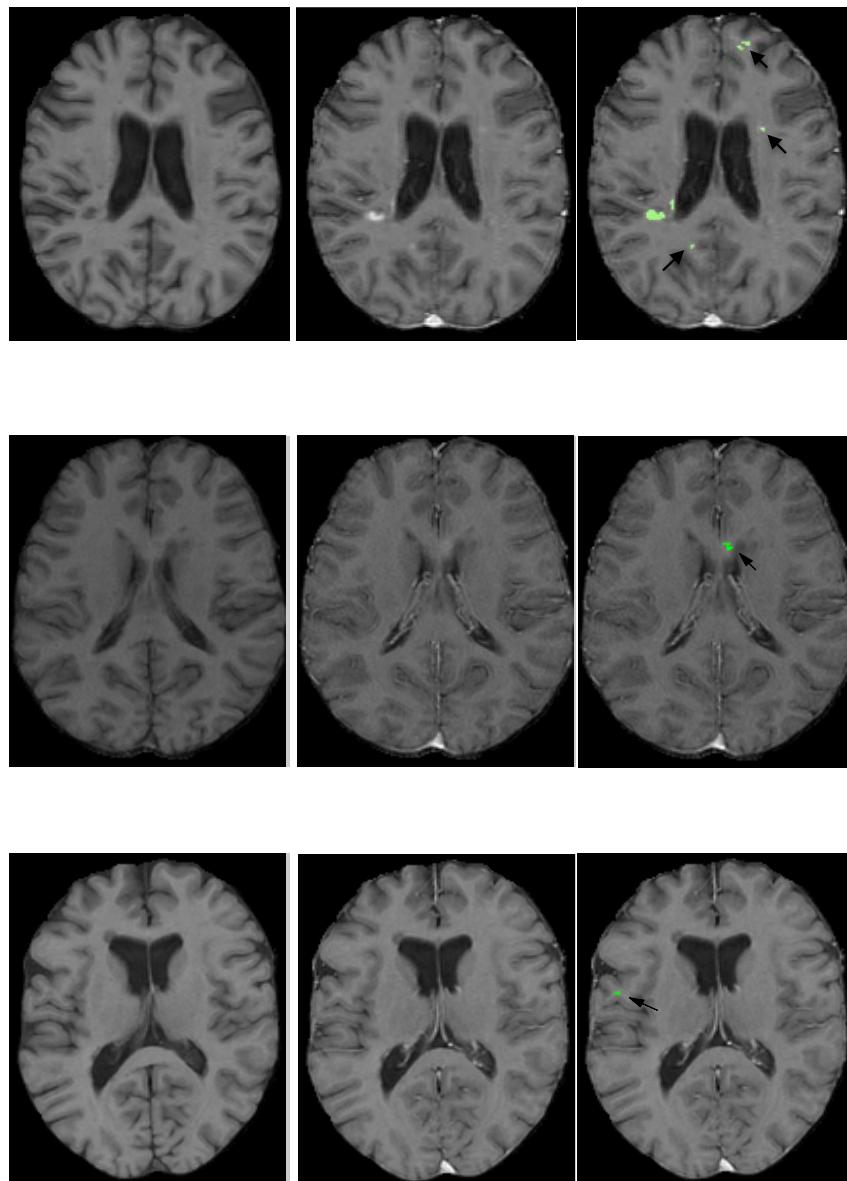


Fig. 7.3 Gad-enhanced lesions. Active MS lesions exposed as hyperintense voxels in the T1-weighted MRI scan after injection of a contrast agent with gadolinium, commonly referred to as Gad lesions. The left column shows axial slices of the pre-contrast T1-weighted MRI scan. The middle column shows the corresponding slices of the post-contrast T1-weighted MRI scan. The right column illustrates the expert-labelled MS lesions (in green) overlayed on the post-contrast T1-weighted MRI scan.

T2 lesions are associated with chronic MS lesions and are commonly exposed by WM hyper-intensities in T2-weighted MRI volumes in conjunction with WM hypo-intensities in T1-weighted MRI volumes, as illustrated in Fig. 7.4. Note that the periventricular lesions exposed in the T2-weighted MRI volume have a clear hyper-intense profile that provides a clear contrast with GM. Unfortunately, the intensity profile is also very similar to the profile of CSF. Such ambiguity justifies the use of a second MRI modality, such as T1-weighted MRI, and healthy tissue probabilities obtained from registration to a brain atlas, which can jointly provide increased discriminability between lesions and CSF. In particular, note how the periventricular lesions exposed in the T1-weighted MRI volume have a slightly brighter profile than CSF. Similar to Gad lesions, the critical regions where registration accuracy is particularly challenging are the WM-GM tissue boundaries found in the cortical regions and the lateral ventricle boundaries.

In short, the two MS lesion detection tasks evaluated in this chapter involve challenging segmentation problems that justify the need for extended information sources, such as a series of MRI sequences and the healthy tissue probability priors obtained via registration to a brain atlas. The next section describes the clinical datasets used to evaluate each method. It is important to note that each context involves a unique dataset, ground truth, and automatic lesion segmentation algorithm. As such, the results obtained from both contexts allow for a *general* comparison of the bottom-line performance of each registration technique evaluated.

The next section describes the clinical dataset used in each context and the corresponding manual labelling protocol employed to generate the ground truth. The subsequent sections illustrate the results obtained in each MS lesion detection contexts. An important and general conclusion obtained from both results is that the proposed technique can provide substantially reduced processing times while leading to slightly improved lesion detection performance.

7.2 Clinical Data

The data used for training and testing each MS lesion detection algorithm was acquired from multi-centre clinical trials with Relapsing Remitting MS (RRMS) patients with varying numbers of lesions located in different areas of the brain. The full proprietary dataset comprises 114 individuals scanned over multiple time points. Each segmentation algorithm

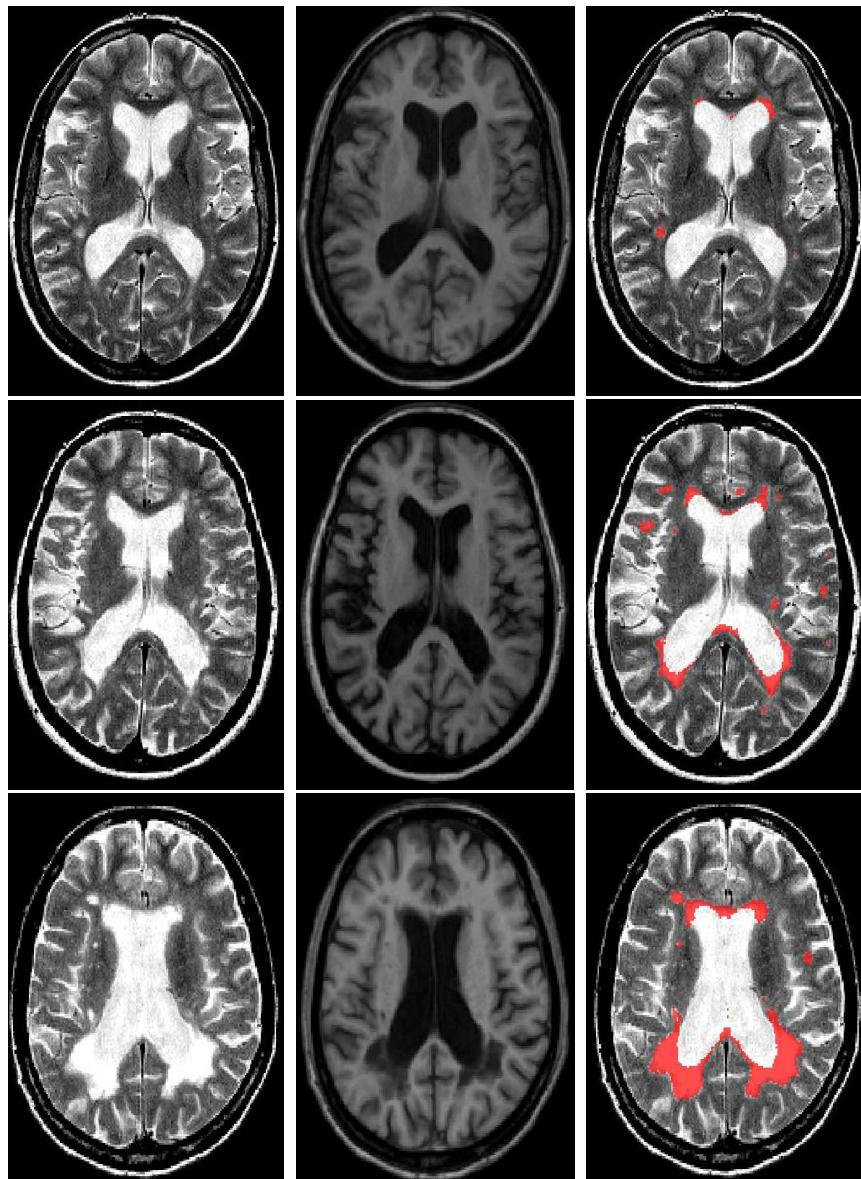


Fig. 7.4 T2 lesions. Chronic MS lesions exposed as hyper-intense WM voxels in the T2-weighted MR volume and hypo-intense WM voxels in the T1-weighted MRI volume, commonly referred to as T2 lesions. The left column shows axial slices of a T2-weighted MRI brain scan from a patient with MS. The middle column shows the corresponding slices of a T1-weighted MRI brain scan. The right column illustrates the expert-labelled MS lesions (in red) overlayed on the T2-weighted MR.

relies on healthy tissue priors obtained by registering the patient’s T1-weighted MRI volume to the ICBM 2009a brain atlas [56] and associated tissue priors. Furthermore, each type of MS lesion involved a strict protocol for establishing the expert ground truth. This section details each of the two protocols involved in the generation of ground truth.

7.2.1 T2 Lesion Ground Truth

The ground truth used for the validation of the T2 lesion detection algorithm is obtained with a semi-automatic process involving an automatic classification of lesions, which is manually verified by removing falsely detected lesions. In particular, the initial automatic procedure involves a Bayesian classification approach based on T1-weighted, T2-weighted and PD-weighted MRI volumes, with a heuristic post-processing stage. The output obtained from such automatic technique is consequently verified by experts so as to remove any false positives.

7.2.2 Gad Lesion Ground Truth

The protocol for defining the ground truth used for validating the automatic detection of Gad lesions involved the independent manual labelling by two trained experts following a common convention, and consequently identifying the labels with a consensus agreement. In particular, the manual labellers followed the convention that a Gad lesion should involve at least a 20% increase in intensity (with reference to the non-contrasted T1-weighted MRI volume) and at least three connected voxels. It is important to note that the expert also has access to all other MRI modalities acquired, which typically include non-contrasted T1-weighted MRI, contrasted (gad-enhanced) T1-weighted MRI, T2-weighted MRI, PD-weighted MRI, and Fluid attenuated inversion recovery (FLAIR) MRI. A silver standard ground truth was then obtained based on the consensus agreement between both experts. If the labellers could not reach an agreement, the case was reviewed by a third highly trained expert.

7.3 Results

This section presents a quantitative analysis of the difference in lesion segmentation performance obtained in the two different brain lesion contexts (i.e. T2 lesion detection and

Gad lesion detection) in terms of different patient to volume registration techniques used to generate healthy tissue priors.

The proposed non-linear registration method, *SymBA*, was compared against two alternative state-of-the-art registration techniques, namely *SyN* [10]¹ and *ANIMAL* [38]. To that end, each segmentation method is re-trained and tested with the set of tissue priors obtained from different registration techniques. The quantitative results are described in the following section.

7.3.1 Gad Lesions

The Gad lesion detection method [81] was trained on a random subset of data containing 86 volumes and tested on 180 volumes. Note that this method relies on the intensities of five MRI sequences (pre- and post-contrast T1-weighted, T2-weighted, PD-weighted and FLAIR) in conjunction with two tissue priors: WM and Partial Volume (PV).

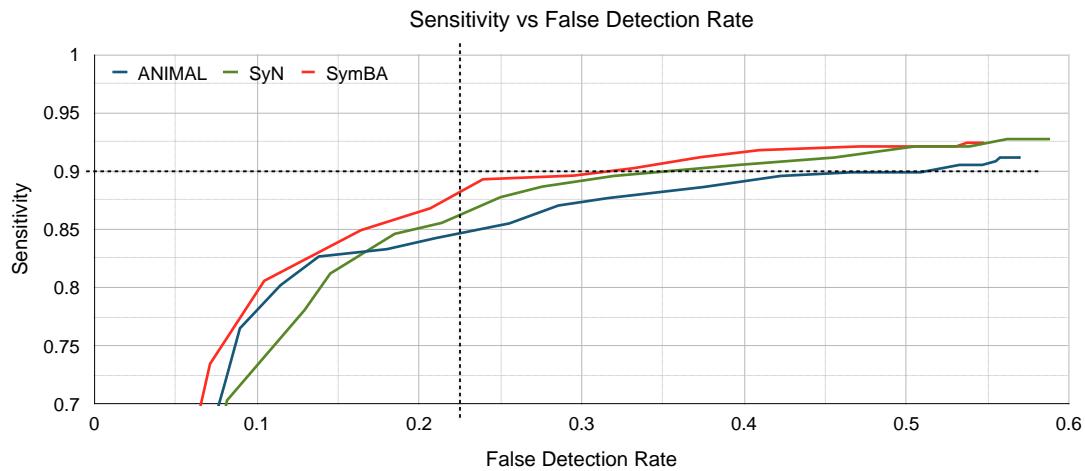


Fig. 7.5 ROC-like curve illustrating the trade-off between sensitivity and FDR for a Gad lesion detection method with varying healthy tissue priors. The dashed lines illustrate working points based on a fixed sensitivity or a fixed FDR.

Fig. 7.5 represents the results obtained in form of ROC-like curves. Each curve shows the trade-off between sensitivity and false detection rates (FDR) for different working points (i.e. thresholds) of the method. The FDR (not to be confused with the false positive

¹The registration technique referred to as *SyN* is part of a general set of neuroimaging software tools known as ANTs, and the two acronyms may be interchanged in the literature.

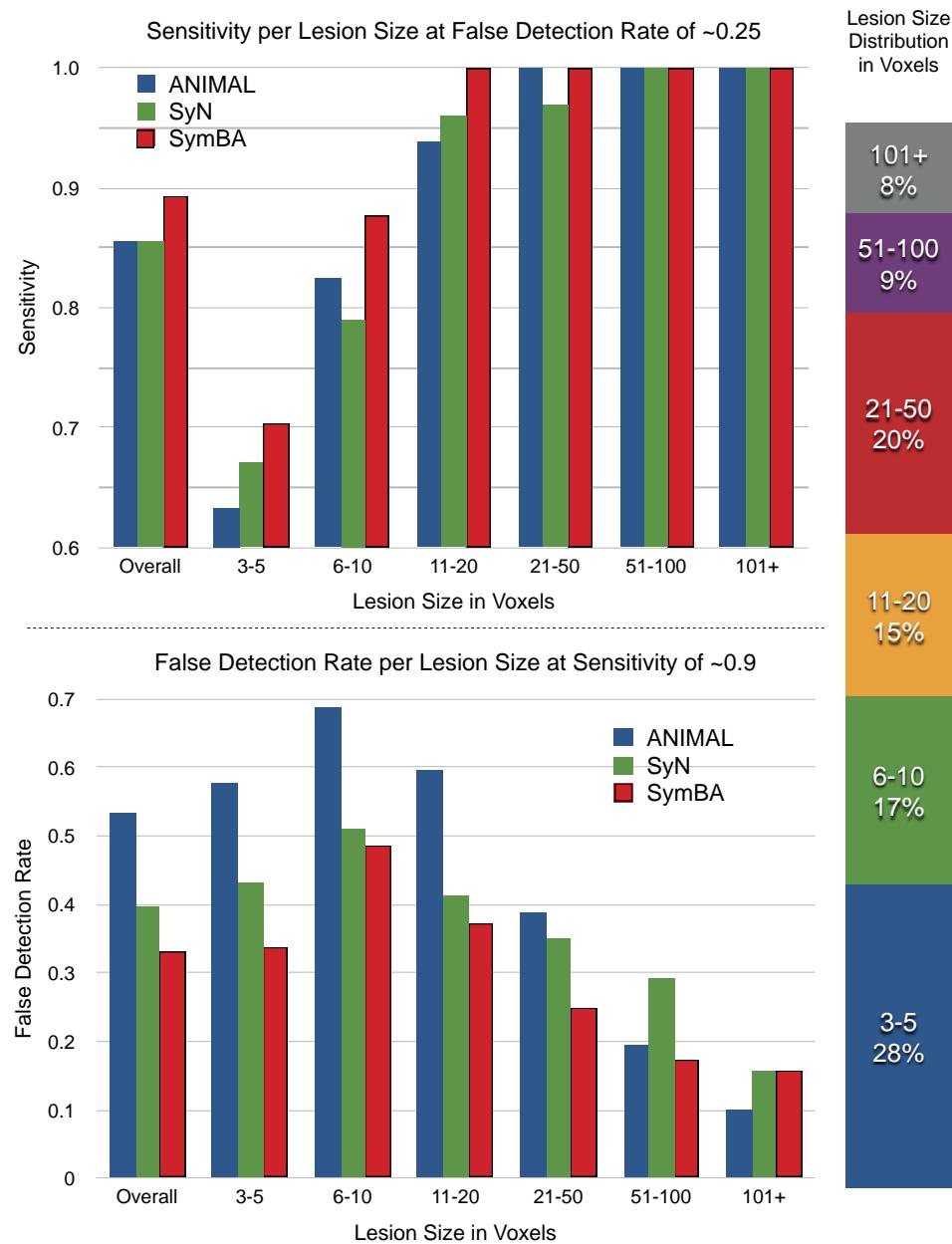


Fig. 7.6 Gad lesion detection performance across lesion size.²

ratio) is defined as the ratio between the number of false positives and the total number of positives. The FDR is commonly used in the context of detection since it is more informative. Note that for almost all working points the proposed registration approach

yields decreased FDR values and increased sensitivity values with relation to the other methods. This is likely due to the fact that SymBA has resulted in more accurate tissue priors containing high WM probabilities at the location of the lesions. Consequently, the corresponding segmentation model has learned to rely more on the priors which has lead to a better detection rate.

The dashed lines in Fig. 7.5 show two operating points, one defined by a fixed sensitivity of 0.9, and another defined by a fixed FDR of 0.25. Both operating points are used to perform a detailed analysis of the detection performance across lesions of different size. Fig. 7.6 shows the distribution of lesions of different size across the testing dataset. The top bar chart in Fig. 7.6 shows the variation of sensitivity across lesion size for a fixed FDR of 0.25. Notice how the proposed approach provides an increase in sensitivity for various lesion sizes leading to an overall increase in sensitivity with relation to other approaches. There is a particularly prominent increase in sensitivity for lesions of size between 3 and 5 voxels, which constitute the largest lesion size category with 28% of the lesions encountered and also constitute the most challenging category of lesions to detect. The bottom chart in Fig. 7.6 shows the variation of FDR for a fixed sensitivity of 0.9. Notice how the proposed approach yields a significantly reduced FDR for almost all lesion size ranges.

7.3.2 T2 Lesions

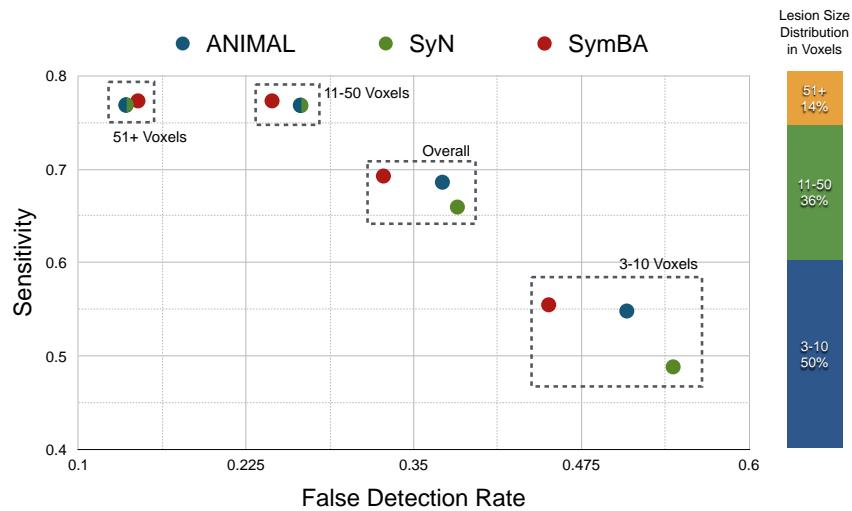


Fig. 7.7 T2 lesion detection performance for different lesion sizes.

The T2 lesion detection method [158] was trained on a random subset of data containing 86 volumes and tested on 28 volumes. Since this method only has one working point, it is not possible to present ROC-like curves. The performance across different ranges of lesions size is shown in Fig. 7.7.

The comparison between registration methods follows a similar trend as the one found for Gad lesions. In particular, the proposed registration approach yields increased overall sensitivity with a reduced FDR. When analyzing the results per lesion size, one can note that the most significant increase in detection performance is found in the smallest lesions with size between 3 and 10 voxels, which constitute the largest lesion size category and the most challenging category of lesions to detect.

7.4 Conclusions

In this chapter, I have presented a set of quantitative results evaluating the variability of lesion detection algorithms with respect to tissue priors from different patient to template registration methods. This constitutes a fundamentally different validation strategy that focuses on the downstream performance of the registration technique, instead of trying to capture the geometric accuracy itself. As such, it represents an indirect measure of the performance of the registration method, yet it directly characterizes the effect of varying registration performance on the specific task of interest.

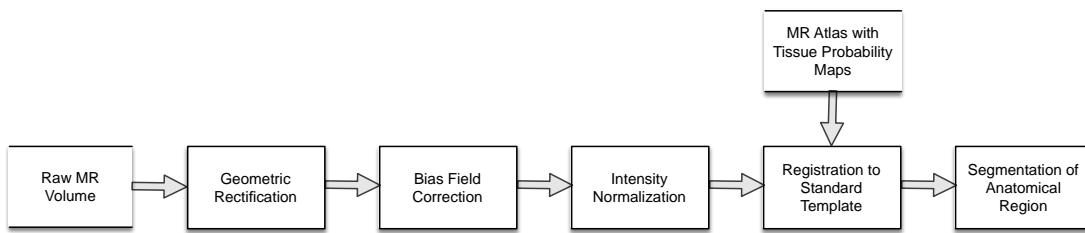


Fig. 7.8 Example of a medical image analysis processing pipeline. Notice how each module is dependent on the output of the previous (left-sided) modules, and any fault from the previous modules can propagate to the final result.

The evaluation strategy employed in this chapter constitutes an important variation in the philosophy for the validation of registration techniques. In particular, instead of isolating the validation to the registration task and attempting to estimate the geometric

accuracy of the registration solution, this evaluation strategy focuses on the bottom-line performance for the particular task at hand and how it is affected by different choices of registration methods. The validation method is essentially characterizing the system-wide application-specific performance with relation to the variation of one module in a downstream chain of dependent modules. Such chain of processing modules is more and more common in modern medical image analysis processing pipelines. Fig. 7.8 shows an example of an image processing pipeline whose goal is to segment an anatomical region from a MRI volume. The processing modules in the pipeline include: 1) the rectification of the raw MRI volume to correct for geometric distortions brought forward by inhomogeneities in the magnetic field, 2) the correction of a low-frequency bias field that breaks the assumption of a spatially constant intensity response, 3) the normalization of voxel intensity values to a “standard” range and distribution, so that their absolute values can be evaluated with a “standard” tissue-intensity model, 4) the registration of the pre-processed MRI volume to a standard atlas which includes a set of tissue probability maps, and 5) the segmentation of the anatomical region of interest based on the pre-processed MRI volume and the deformed tissue probability maps obtained from the previous registration. It is clear that the performance of any segmentation algorithm (found in the last module) will be a function of the performance of all previous processing modules, as well as the quality of the template and its tissue probability maps. Furthermore, since a majority of those modules continue to be active research topics, it would be a critical mistake to automatically assume that any off-the-shelf solution would be sufficiently good for the task at hand.

One can also consider the proposed evaluation strategy as an approach for characterizing the *robustness* of the segmentation algorithm to performance variations from the other modules. For example, in the context of brain lesion detection, one could measure the sensitivity of the detection technique to induced variations on the patient to template transformation. Such measure would roughly characterize the amount of performance that is lost or gained as a function of registration accuracy and the information value obtained from the tissue probability maps. This type of analysis is fundamentally important since it is not always clear what specific components of the registration performance are critical to the bottom-line performance. As such, registration techniques that yield an improved registration accuracy averaged over the spatial domain, may actually involve a poor performance in a critical region for the task at hand and consequently lead to a decreased bottom-line performance.

In summary, the validation strategy found in this chapter has been proposed so as to emphasize the importance of characterizing the effect of a registration technique on subsequent modules. The experimental results shown demonstrate that the proposed technique can lead to improved brain lesion detection performance with two different segmentation approaches, while also involving significantly reduced processing times. This particular task of brain lesions detection can be found in the context of clinical trials for treatments where the MRI volumes are analyzed for relevant biomarkers to see if the treatment has a statistically significant effect on the population. Such clinical trials generally involve thousands of scans at different timepoints from hundreds of patients and involve multiple types of analysis on the images. Hence, any significant improvement in computational efficiency is highly valuable for both reducing the financial cost of extensive computations and reducing the turnaround time necessary for evaluating results at the scale of the clinical trial.

Chapter 8

Conclusions and Future Directions

Image registration is a fundamental problem encountered in a variety of medical image analysis tasks. The diversity of medical imaging modalities encountered throughout all kinds of clinical contexts has led to the development and proposal of many different techniques with an increasing trend toward application-specific algorithms. Two of the most recurrent limitations of current techniques involve the high processing times involved in non-rigid contexts and the limited robustness towards multi-modal contexts with widely different image formation models.

The work presented in this thesis describes a general registration framework which addresses such limitations and aims to achieve a significant reduction in processing times, as well as a significant improvement in robustness in challenging multi-modal scenarios. In particular, the proposed framework focuses on the use of image gradient orientations at sparsely selected voxel locations. A major motivation behind such approach lies in using a primitive geometric descriptor for assessing the similarity of two mapped image locations, so as to somewhat decouple the notion of image similarity from the particular intensity profile which may include imaging artefacts from the particular imaging device. Such decoupling of the image similarity metric from the voxel intensities is particularly relevant in multi-modal contexts (e.g. MRI to US) where a global relationship between corresponding voxel intensities simply does not hold. A second motivation lies in focusing or restricting the evaluation of image similarity to regions of interest that correspond to underlying boundaries in the scene. These are the major principles behind the development of the two proposed and closely related techniques.

The first proposed technique addresses the rigid registration of challenging multi-modal contexts and is characterized as the maximization of image gradient orientation alignment evaluated on prudently selected locations in the fixed image domain. For the purposes of defining a consistent multi-scale voxel selection scheme, this work provided an analysis of gradient orientation uncertainty brought forward by an additive Gaussian noise model. In particular, the task of evaluating the gradient orientation of the underlying scene (i.e. undegraded image) is defined as a Bayesian inference problem, where the computed image gradient magnitude and orientation constitute “observations” which shape the posterior distribution of the “true” underlying gradient orientation. Such analysis demonstrated that the gradient magnitude is an indicator of the orientation variance and that normalizing the gradient magnitude with respect to the corresponding derivative kernel leads to a “scale-independent” indicator of gradient orientation uncertainty. Hence, the proposed technique relies on the manual definition of a single gradient magnitude threshold for the finest image scale which can be mapped to corresponding thresholds at coarser image scales.

The proposed rigid registration technique also involves multiple evaluations of the image gradient of a transformed moving image, which can lead to a significant computational overhead when evaluated directly from the transformed (i.e. resampled) moving image intensities. In order to minimize such computational overhead, the work presented in this thesis demonstrated that the pullback differential can be used to evaluate the gradient of the transformed moving image with computational efficiency. Specifically, the transformed image gradient is computed by simply multiplying the original image gradient found at the transformed location by the transpose of the Jacobian of the transformation function, which is already known and corresponds to the transpose of the rotation matrix.

The second proposed technique addresses non-rigid registration contexts and involves the maximization of image gradient orientation alignment as well as the minimization of the distance to the nearest inferred boundary in a symmetric fashion. As such, it relies on the explicit inference of boundaries with the use of an edge detection technique, which defines a reduced set of locations over which the local similarity metric is evaluated and allows the evaluation of a distance map where each voxel’s value corresponds to the distance to the nearest boundary location.

The proposed non-rigid registration technique also relies on a symmetric evaluation of the local similarity metric at locations identified in the fixed image domain (using the forward transform) as well as locations identified in the moving image domain (using the

backward transform). In order to allow the evaluation of such an energy function, the method adopted a diffeomorphic transformation model characterized as the integration of a velocity field. Furthermore, in order to evaluate the gradient orientations of a transformed image, the work presented in this thesis developed a pushforward technique which reduces the computational operations necessary to evaluate it and also shares many of the operations involved in the evaluation of the transformed location. Finally, this work also presented the necessary derivations and expressions required to evaluate the gradient of the proposed energy function with respect to the transformation parameters (i.e. the voxel values of the velocity field image).

The proposed techniques were then shown to provide important advantages in multiple challenging clinical contexts. In particular, the rigid registration technique led to substantial improvements in registration performance in the context of IGNS with tracked iUS, where the proposed technique can successfully register a pre-operative MRI to a iUS with sub-second processing times. Such an accomplishment leads to a seamless integration in the surgical procedure where the clinician can immediately obtain an updated MRI volume after an US acquisition is performed. The advantages brought forward by the proposed framework were then shown to generalize to multiple clinical scenarios. In particular, the results presented in this thesis have demonstrated that the proposed technique can provide important advantages in the contexts of the rigid registration of CT and MRI brain volumes, the rigid registration of pre-operative MRI to iUS for IGNS, the non-rigid registration inter-patient MRI brain volumes for the automatic anatomical labelling of brain structures, and the non-rigid registration of patient to template MRI brain volumes for the generation of healthy soft tissue probability maps. A brief summary of the advantages and contributions provided in each context can be found below.

In the context of the rigid registration of CT and MRI brain volumes from a standard dataset, the reported experiments exposed the importance of sub-voxel (i.e. off the voxel grid) location sampling and of using a highly selective gradient orientation alignment metric for obtaining a significant improvement in registration accuracy. More importantly, a performance comparison with an alternative state-of-the-art voxel selection strategy (used with a NMI similarity metric) demonstrated that the proposed technique can achieve sub-millimeter registration accuracy with drastically reduced sampling rates. In particular, using only 0.05% of the voxels from the fixed image already yielded a median registration error inferior to one millimeter.

In the context of IGNS, the results presented in this thesis illustrated an exceptional performance in the rigid registration of pre-operative MRI to iUS, where the proposed technique (as implemented in a GPU) was shown to successfully register 14 real clinical cases with sub-second processing times. This constitutes an unprecedented functionality which is now a part of a neurosurgical navigation platform employed in ongoing studies at the Montreal Neurological Institute.

In the context of the non-rigid registration of inter-patient MRI brain volumes, the proposed framework was evaluated on a publicly available and commonly used dataset against multiple state-of-the-art techniques and was shown to provide closely similar accuracy to the top ranking methods but with drastically reduced processing times. In particular, a leading alternative state-of-the-art registration technique yielded a compound mean target overlap of 51.63% with an average processing time of 38 minutes, while the proposed technique achieved a compound mean target overlap of 51.65% with an average processing time of 6 minutes and 30 seconds, and a compound mean target overlap of 50.69% with an average processing time of 1 minute and 30 seconds with an alternative parametrization.

Finally, the proposed framework was also evaluated in the context of patient to template MRI brain volume registration for the purposes of generating healthy tissue probabilities for each voxel in the patient's volume. Such healthy tissue probability maps constitute a critical input for automatic brain lesion detection algorithms. The results presented in this thesis demonstrated that the proposed technique led to a slightly improved detection performance in two distinct MS lesions detection tasks and with drastically reduced processing times in comparison with alternative state-of-the-art registration algorithms. Such improvements in performance constitute an important advantage for high data throughput pipelines such as the analysis of large clinical trial datasets.

8.1 Discussion and Future Directions

The proposed registration techniques were characterized by a set of algorithmic parameters and components that had to be suitably defined for the context at hand. Some of the most critical components of the registration techniques are those related to the inference of underlying boundaries (e.g. gradient magnitude threshold) as well as those related to defining and shaping the space of plausible transformation functions. Furthermore, the experimental results in this thesis adopted a particular set of validation strategies for the

evaluation of registration performance with a unique set of advantages and disadvantages. The following sections discusses these and other points and outlines some possible strategies that can be pursued to as to overcome some of the related limitations.

8.1.1 Edge Detection

The two proposed techniques rely on a prudent, yet likely sub-optimal, inference of underlying boundaries which consequently define the energy function to be optimized and necessarily characterize the final performance of the registration methods. It is important to note that the sensitivity of the registration performance to the accuracy of the inferred boundaries is directly related to the degrees of freedom found in the transformation model. For example, in the rigid registration context, the proposed technique relied on a gradient magnitude threshold to essentially ignore a majority of locations that did not correspond to an underlying anatomical boundary. Hence, even though the selected locations did not provide a pixel-compact localization of an underlying boundary they were indeed more than sufficient to jointly identify an accurate global maxima in the space of rigid transformation. Alternatively, in the non-rigid registration context, the proposed technique adopted an explicit dependence to a pixel-compact edge detection technique, such as the Canny edge detection method, which provided the increased local discriminability required to resolve the diffeomorphic transformation model.

It is also important to note that the relevance of the performance of the edge detection technique on the registration method cannot be strictly evaluated in terms of its robustness and accuracy towards identifying anatomical boundary in a single given image. In particular, one should note that in multi-modal registration contexts the edge detection technique should ideally identify the anatomical boundaries that are exposed in *both* modalities and ignore all other exposed boundaries since they cannot serve to resolve the spatial mapping between volumes. For example, the multi-modal registration scenarios evaluated in this work relied on the assumption that the boundaries exposed in one modality (e.g. US, CT) could be assumed to have a corresponding exposed boundary in a second modality (e.g. MRI), and this allowed the definition of an asymmetric voxel selection scheme which focuses on boundary locations identified in one modality. However, other multi-modal scenarios surely involve a more complicated relationship between boundaries exposed in each modality. This immediately brings forward a challenge that goes well-beyond any state-of-the-art

edge detection technique, since it presumes additional higher level prior knowledge, such as knowledge of the image formation model of both modalities. An appealing strategy to identify the image boundaries that are expected to have a correspondence in a second modality could aim to learn such relationship with a set of example cases with known ground truth. One could then exploit such knowledge in a two-stage edge detection strategy, where the first stage corresponds to a conventional edge detection technique that identifies all exposed image boundaries and the second stage corresponds to a pruning technique that rejects all boundary locations that are not likely to have a correspondence in the second modality.

In the interest of generality and of removing the dependency to a separate edge detection stage, it would also be of interest to consider dynamically coupling the inference of underlying anatomical boundaries with the inference of the spatial mapping between images. In other words, an extension of this work would be to consider a *joint* inference task where underlying anatomical boundaries and spatial correspondence are jointly modelled. This effectively constitutes a joint segmentation and registration problem with the distinction that segmentation is focused on anatomical boundaries as opposed to anatomical regions or tissue types.

8.1.2 Regularization and Parameterization of the Transformation Space

Section 6.2.1 presented empirical results showing how different spatial and temporal discretizations of the diffeomorphic transformation model affected the registration accuracy. As such, it reflects the registration performance with varying “hard” restrictions on the space of plausible transformation functions. A similar experiment based on a “soft” or continuous restrictions on the space of plausible transformation functions could evaluate the variation in registration performance with a varying weight assigned to a regularization term (e.g. bending energy).

The main goal of such hard or soft restrictions on the space of transformation functions is to overcome the ill-posedness of registration problems, particularly in contexts involving non-rigid deformations where local displacements are independent between distant regions. Yet, it remains unclear how to define what is the precise degree and type of regularization that is best suited for a specific clinical context.

One way of viewing such issue is in terms of *parameter tuning*. In particular, the proposed non-rigid registration techniques involved a series of configuration parameters

that must be set to “reasonable” values so as to obtain the performance gains reported in this work. This thesis has included a detailed quantitative analysis of performance in terms of registration parameters. As such, it constitutes a valuable guide for choosing suitable parameter values. Nonetheless, it would be of significant interest to pursue a fully automated algorithm for identifying the optimal set of parameter values based on a training set with known ground truth. One can think of this as an optimization problem where the goal is to maximize registration accuracy across training cases while minimizing some notion of deformation energy (e.g. bending energy). The benefit of such a strategy would be that non-technical users could successfully use and evaluate a technique in their particular task without having to understand the algorithmic functionality of the parameters.

8.1.3 Context Learning

The two fundamentally important components of an image registration technique that must be defined for addressing a specific context are the image similarity term and the space and plausibility of geometric transformations mapping one coordinate space to the other. Typically, it is the “designer” (i.e. graduate student) that arbitrarily defines such components by alluding to basic principles or his or her higher level intuition as to what make a reasonable approach for the task at hand. However, as it has been pointed out in the previous sections, it would be of interest to seek a higher level method which itself resolves the best suited registration approach for a specific context by learning the specific image and deformation properties from a set of example cases with known ground truth.

It is important to note that even a higher level learning method for registration inevitably involves an arbitrary space of plausible registration methods that can be learned or deduced. In other words, it is rather unrealistic to envision a top level learning method which evaluates the utility of all possible registration methods and identifies the optimal approach. Instead, one can simply expect the resolution of more limited questions such as:

- Is the image similarity term location dependent (i.e. does the image similarity vary significantly in different matched regions)?
- What is the maximal weight than can be assigned to a regularization term that maintains the global maxima of the energy term (involving the similarity and regularization term) within an acceptable error?

- Which image similarity measure (from a finite predefined set) has a global maxima with best accuracy?
- Which image similarity provides the most consistent and wide aperture range?

Nonetheless, even if the scope of such questions and corresponding learning methods can be considered somewhat limited they would still provide a valuable tool for facilitating the generalization and parametrization of a set of registration techniques to a specific task at hand.

8.1.4 Validation

Many medical image analysis problems face the complicated task of defining a suitable validation strategy and the research community is continuously discussing the limitations of the strategies in use. There are essentially two major options for validation in the context of medical image registration: rely on synthetically generated (simulated) cases with known ground truth deformations, or rely on an “expert” for defining a gold standard on a set of real clinical cases.

The first approach (simulation of cases) provides a perfectly accurate ground truth that is defined throughout the entire image domain. As such, it allows for the evaluation of multiple types of geometrical measures regarding registration accuracy. However, one can immediately put into question the ability of such simulated cases in reflecting all the subtle variabilities found in real clinical cases. This concern can easily be exemplified in medical image modalities, such as MRI and US, that are afflicted by complex image artefacts such as: geometric distortion brought forward by non-linearities in the magnetic field, a bias field that leads to a non-homogeneous image intensity response, motion artefacts, US speckle, shadow and direction-dependent imaging characteristics, among many others.

The second approach (expert gold standard) provides an imperfect ground truth that is typically limited to a relatively sparse set of annotations. Hence, even though it indeed reflects all the challenges and subtle variabilities found in real clinical cases, it provides a measure of performance that is ultimately bounded by the expert’s ability to consistently and accurately define a gold standard.

This thesis adopted the second approach, and relied on validation approaches based on real clinical cases and evaluated with an expert defined gold standard. As such, the

challenges and subtleties embedded in a set of real clinical cases were considered to be more relevant, in practical terms, than the pursuit of hyper accurate registration performance (e.g. sub-voxel accuracy) on synthetic cases. Furthermore, the experimental setups relied on publicly available and commonly used datasets, and thus allow the direct and fair comparison of reported results with results presented in previous work.

Nonetheless, there are important limitations to such validation approaches. In particular, Section 4.5 illustrated that the use of expert defined homologous landmarks between MRI and US volumes came with a significant degree of variability between experts, and as such provided an imperfect measure of registration accuracy. Furthermore, the distribution of such landmarks was arbitrarily established by each expert's notion of visual and anatomical correspondence, and such distribution does not necessarily characterize the registration accuracy consistently throughout the clinical region of interest. However, in that particular context, the dataset used represents the sole publicly available dataset of real clinical cases of brain tumour resection surgeries and likely constitutes the most relevant validation strategy presently available for the registration between such modalities.

In short, the validation strategies adopted in this work were justified by the fact that they directly characterize registration performance with relation to the concrete clinical task of interest.

8.1.5 Other Extensions

The work presented in this thesis has led to highly efficient rigid and non-rigid image registration algorithms with direct applications in medical contexts. Nonetheless, the proposed framework can be further extended by either continuing the effort of further reducing processing times or by extending it for suitable use in alternative contexts.

For the purposes of further reduction in processing times, a natural extension of the proposed non-rigid registration technique would be to develop a GPU-based implementation. This particular exercise was developed with the proposed rigid registration technique and led to substantial reductions in processing times. Since the proposed non-rigid registration technique largely follows a similar algorithmic structure, one should expect similar reductions in processing times. In particular, a GPU-based implementation should achieve the same accuracy but with processing times in the order of seconds or less.

Similarly, it would be of interest to pursue optimization variants with a *fixed* processing

time constraint that allows for continuous feedback to an end-user at specific times intervals, at the possible expense of a small degradation in registration accuracy. This functionality may be particularly relevant in the context of active real-time imaging contexts, such as US imaging, where the clinician/user can continuously adapt the position of the imaging device (e.g. US probe) so as to optimally expose the anatomical region of interest. For example, one can envision a near real-time image analysis process based on registration that continuously provides an updated visualization of pre-segmented structures after an inferred deformation, and where the clinician can adapt the image device in tandem with this continuous image registration and analysis process.

The reduced processing times of the proposed non-rigid registration strategy may also play a very important role in multi-atlas label fusion strategies, where the segmentation labels of a given image are obtained by registering the image against a set of pre-segmented templates and consequently fusing the spatially mapped template labels. It is clear that the main computational cost of such approaches lies in the series of non-rigid registrations from patient to template, and thus an accurate yet computationally efficient registration technique can lead to drastically reduced processing times. In particular, it would be of interest to characterize the trade-off in segmentation performance obtained with a given budget of computational time per registration, where one setup adopts a highly accurate but costly registration approach that allows the use of only a small number of templates, and an alternative setup adopts a slightly less accurate but computationally inexpensive registration approach that allows the use of a large number of reference atlases.

The work presented here focused on the use of image gradient orientations as features of interest, where image gradient orientations are primitive geometric descriptors which relate to the normal direction of an underlying boundary. Since the local feature is geometric and decoupled from the specific image intensities, it can also be compared with purely synthetic 3D objects, such as surface meshes. As such, it would be of interest to adopt the proposed registration techniques in registration contexts involving a mixture of 3D image volumes and 3D surfaces, such as the ones obtained from manually segmented data. Such an approach can be highly relevant in the scenario when one relies on a patient to template registration for segmenting an anatomical region of interest. Instead of relying on a purely volumetric registration (between two 3D volumes), one could directly register a template 3D surface of the anatomical region of interest to the patient's volume.

Finally, the proposed diffeomorphic registration technique focused on a computationally

efficient image similarity term, but did not devote an effort to ensure that the resulting transformation accurately satisfies the diffeomorphic properties of smoothness and invertibility. It would be of interest to explicitly enforce such properties by adopting a regularization term that rewards smooth transformations (e.g. the L^2 norm of a second-order differential operator of the velocity field) and adjust the number of time steps such that the forward and backward transformation are consistent within a predefined accuracy.

Automated medical image analysis holds the promise of providing consistent, unbiased, and accurate diagnosis with sophisticated computational techniques which require no human supervision. As such, it is an exciting and highly active field which I believe will eventually become a critical component of health care; where it may provide improvements not only in patient prognosis but perhaps more importantly lead to significant reductions in the financial cost associated with providing such care. The accuracy and computational efficiency of medical image registration techniques are, without a doubt, an important part of such promise and must therefore continue to be improved. I hope that the methods and results shown in this thesis will provide a step forward in fulfilling such vision.

Appendix A

Mathematical Glossary

Symbols

The following table summarizes the most commonly used mathematical symbols in this thesis.

Ω_f	fixed image spatial domain	$\Omega_f = \mathbb{R}^d$ where $d \in \{2, 3\}$
Ω_m	moving image spatial domain	$\Omega_m = \mathbb{R}^d$ where $d \in \{2, 3\}$
\mathbf{x}_f	location in fixed image domain	$\mathbf{x}_f \in \Omega_f$
\mathbf{x}_m	location in moving image domain	$\mathbf{x}_m \in \Omega_m$
t	time index	$t \in \{1, \dots, \tau\}$
τ	number of time steps	
Δt	time step size	$\Delta t = \frac{1}{\tau}$
I_f	fixed image	$I_f : \Omega_f \rightarrow \mathbb{R}^n$ (e.g. $n = 1$ for a scalar image)
I_m	moving image	$I_m : \Omega_m \rightarrow * \mathbb{R}^n$ (e.g. $n = 1$ for a scalar image)

\mathbf{T}_f	forward mapping transformation function	$\mathbf{T}_f : \Omega_f \rightarrow \Omega_m$
\mathbf{T}_m	backward mapping transformation function	$\mathbf{T}_m : \Omega_m \rightarrow \Omega_f$
∇I	image gradient	$\nabla I = (\frac{\partial I}{\partial x_1}, \dots, \frac{\partial I}{\partial x_D})$
\mathbf{v}_u	velocity field at time index u	$\mathbf{v} = \frac{d\mathbf{T}_f}{dt} \Big _{t=u}$
$\phi_{f,\tau}$	forward integrated velocity field	$\phi_{f,t}(\mathbf{x}_f) = \mathbf{x}_f + \sum_{t=1}^{\tau} \mathbf{v}_t(\phi_{f,t-1}) \Delta t$
$J_f(\mathbf{x})$	Jacobian of f with respect to x	$J_f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} \dots \frac{\partial f_1}{\partial x_D} \\ \vdots \ddots \vdots \\ \frac{\partial f_D}{\partial x_1} \dots \frac{\partial f_D}{\partial x_D} \end{bmatrix}$
○	composition operator	

References

- [1] P. Aljabar, R. A. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *NeuroImage*, 46(3):726–738, July 2009.
- [2] R. Alterovitz, K. Goldberg, J. Pouliot, I.-C. J. Hsu, Y. Kim, S. M. Noworolski, and J. Kurhanewicz. Registration of MR prostate images with biomechanical modeling and nonlinear parameter estimation. *Medical Physics*, 33(2):446–454, Jan. 2006.
- [3] N. A. Álvarez, J. M. Sanchiz, J. Badenas, F. Pla, and G. Casañ. Contour-Based Image Registration Using Mutual Information. In *Pattern Recognition and Image Analysis*, pages 227–234. Springer Berlin Heidelberg, Berlin, Heidelberg, Jan. 2005.
- [4] T. Arbel, X. Morandi, R. M. Comeau, and D. L. Collins. Automatic Non-linear MRI-Ultrasound Registration for the Correction of Intra-operative Brain Deformations. In *Proceedings of the Conference for Medical Image Computing and Computer Assisted Intervention - MICCAI 2001*, pages 913–922. Springer-Verlag, Utrecht, Netherlands, Oct. 2001.
- [5] V. I. Arnold and B. A. Khesin. Topological methods in hydrodynamics. *Topological methods in hydrodynamics*, Jan. 1998.
- [6] Asgeir S Jakola, Geirmund Unsgård, and Ole Solheim. Quality of life in patients with intracranial gliomas: the impact of modern image-guided surgery. *dx.doi.org*, 114(6):1622–1630, June 2011.
- [7] J. Ashburner and K. J. Friston. Unified segmentation. *NeuroImage*, 26(3):839–851, 2005.
- [8] B. Avants, C. Anderson, M. Grossman, and J. C. Gee. Spatiotemporal Normalization for Longitudinal Analysis of Gray Matter Atrophy in Frontotemporal Dementia. pages 303–310, 2007.
- [9] B. Avants, P. A. Cook, C. McMillan, M. Grossman, N. J. Tustison, Y. Zheng, and J. C. Gee. Sparse unbiased analysis of anatomical variance in longitudinal imaging.

- In *MICCAI'10: Proceedings of the 13th international conference on Medical image computing and computer-assisted intervention: Part I*. Springer-Verlag, Sept. 2010.
- [10] 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. *Medical Image Analysis*, 12(1):26–41, 2008.
 - [11] B. B. Avants, N. J. Tustison, J. Wu, P. A. Cook, and J. C. Gee. An open source multivariate framework for n-tissue segmentation with evaluation on public data. *Neuroinformatics*, 9(4):381–400, Dec. 2011.
 - [12] B. B. Avants, P. Yushkevich, J. Pluta, and D. Minkoff. The optimal template effect in hippocampus studies of diseased populations. *NeuroImage*, 2010.
 - [13] R. Bajcsy, R. Lieberson, and M. Reivich. A Computerized System for the Elastic Matching of Deformed Radiographic Images to Idealized Atlas Images. *Journal of Computer Assisted Tomography*, 7(4):618, Aug. 1983.
 - [14] S. K. Balci, P. Golland, and W. Wells. Non-rigid groupwise registration using B-spline deformation model. *Open source and open data for ...*, 2007.
 - [15] D. I. Barnea and H. F. Silverman. A Class of Algorithms for Fast Digital Image Registration. (2):179–186, 1972.
 - [16] 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*, 61(2):139–157, 2005.
 - [17] E. M. Berntsen, S. Gulati, O. Solheim, and K. A. Kvistad. ... magnetic resonance imaging and diffusion tensor tractography incorporated into an intraoperative 3-dimensional ultrasound-based neuronavigation system: impact, 2010.
 - [18] R. Bhagalia, J. A. Fessler, and B. Kim. Accelerated Nonrigid Intensity-Based Image Registration Using Importance Sampling. *IEEE TMI*, 28(8):1208–1216, Aug. 2009.
 - [19] F. L. Bookstein. Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.
 - [20] R. Brooks. *Efficient and Reliable Methods for Direct Parameterized Image Registration*. PhD thesis, McGill University, 2008.
 - [21] R. Brooks and T. Arbel. The importance of scale when selecting pixels for image registration. In *Computer and Robot Vision, 2007. CRV '07. Fourth Canadian Conference on*, pages 235–242. IEEE, 2007.

- [22] R. Brooks, D. L. Collins, X. Morandi, and T. Arbel. Deformable Ultrasound Registration without Reconstruction. In *MICCAI '08: Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention, Part II*, pages 1023–1031. Springer-Verlag, Sept. 2008.
- [23] L. G. Brown. A Survey of Image Registration Techniques. *ACM Computing Surveys*, 24(4):325–376, Dec. 1992.
- [24] M. Brown and D. G. Lowe. Automatic Panoramic Image Stitching using Invariant Features. *International Journal of Computer Vision*, 74(1):59–73, 2007.
- [25] Y. Cao, M. I. Miller, R. L. Winslow, and L. Younes. Large deformation diffeomorphic metric mapping of vector fields. *IEEE TMI*, 24(9):1216–1230, Sept. 2005.
- [26] V. S. Caviness Jr, J. Meyer, N. Makris, and D. N. Kennedy. MRI-based topographic parcellation of human neocortex: an anatomically specified method with estimate of reliability. *Journal of Cognitive Neuroscience*, 8(6):566–587, 1996.
- [27] H. Chang and J. M. Fitzpatrick. A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities. *Medical Imaging, IEEE Transactions on*, 11(3):319–329, Sept. 1992.
- [28] S. D. Chang, W. Main, D. P. Martin, I. C. Gibbs, and M. P. Heilbrun. An Analysis of the Accuracy of the CyberKnife: A Robotic Frameless Stereotactic Radiosurgical System. *Neurosurgery*, 52(1):140, Jan. 2003.
- [29] C. Chefd'Hotel, G. Hermosillo, and O. Faugeras. *A variational approach to multi-modal image matching*. IEEE, 2001.
- [30] H.-M. Chen and P. K. Varshney. Mutual information-based CT-MR brain image registration using generalized partial volume joint histogram estimation. *IEEE TMI*, 22(9):1111–1119, 2003.
- [31] T. K. Chen, A. D. Thurston, and R. E. Ellis. A real-time freehand ultrasound calibration system with automatic accuracy feedback and control. *Ultrasound in medicine & ..., 2009*.
- [32] G. E. Christensen and H. Johnson. Consistent Image Registration. *IEEE Transactions on Medical Imaging*, 20(7):568–582, 2001.
- [33] G. E. Christensen, R. D. Rabbitt, and M. I. Miller. Deformable templates using large deformation kinematics. *Image Processing, IEEE Transactions on*, 5(10):1435–1447, 1996.

- [34] O. Clatz, H. Delingette, I. Talos, A. J. Golby, R. Kikinis, F. A. Jolesz, N. Ayache, and S. K. Warfield. Robust Nonrigid Registration to Capture Brain Shift From Intraoperative MRI. *IEEE Transactions on Medical Imaging*, 24(11):1417–1427, 2005.
- [35] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. In *Information processing in medical imaging*, pages 263–274. Kluwer Academic Publishers, 1995.
- [36] D. L. Collins and A. C. Evans. ANIMAL: Validation and applications of nonlinear registration-based segmentation. *International Journal of Pattern Recognition and Artificial Intelligence*, 11(8):1271–1294, 1997.
- [37] D. L. Collins, C. J. Holmes, T. M. Peters, and A. C. Evans. Automatic 3-D model-based neuroanatomical segmentation. *Human Brain Mapping*, 3(3):190–208, 1995.
- [38] D. L. Collins, A. P. Zijdenbos, W. F. Baaré, and A. C. Evans. ANIMAL+ INSECT: improved cortical structure segmentation. In *Information processing in medical imaging*, pages 210–223. Springer, 1999.
- [39] M. M. Coselmon, J. M. Balter, D. L. McShan, and M. L. Kessler. Mutual information based CT registration of the lung at exhale and inhale breathing states using thin-plate splines. *Medical Physics*, 31(11):2942–2948, 2004.
- [40] P. Coupé, P. Hellier, X. Morandi, and C. Barillot. 3D rigid registration of intra-operative ultrasound and preoperative MR brain images based on hyperechogenic structures. *Journal of Biomedical Imaging*, 2012.
- [41] Y. Dai, J. Tian, D. Dong, and G. Yan. Real-time visualized freehand 3D ultrasound reconstruction based on GPU. . . . in *Biomedicine*, 2010.
- [42] M. H. Davis, A. Khotanzad, D. P. Flamig, and S. E. Harms. A Physics-Based Co-ordinate Transformation for 3D Image Matching. *IEEE Transactions on Medical Imaging*, 16(3):317–328, 1997.
- [43] D. De Nigris, T. Arbel, and D. L. Collins. SymBA: Diffeomorphic Registration Based on Gradient Orientation Alignment and Boundary Proximity of Sparsely Selected Voxels. In *6th International Workshop on Biomedical Image Registration*.
- [44] D. De Nigris, D. Collins, and T. Arbel. Fast and robust registration based on gradient orientations: case study matching intra-operative ultrasound to pre-operative MRI in neurosurgery. pages 125–134. Springer, 2012.

- [45] D. De Nigris, D. Collins, and T. Arbel. Multi-modal image registration based on gradient orientations of minimal uncertainty. *Medical Imaging, IEEE Transactions on*, 31(12):2343–2354, Dec 2012.
- [46] D. De Nigris, D. L. Collins, and T. Arbel. Deformable Registration of Chest CT Scans with Adaptive Local Mutual Information. In *Grand Challenges in Medical Image Analysis*, 2010.
- [47] D. De Nigris, D. L. Collins, and T. Arbel. Fast rigid registration of pre-operative magnetic resonance images to intra-operative ultrasound for neurosurgery based on high confidence gradient orientations. *International Journal of Computer Assisted Radiology and Surgery*, 8(4):649–661, 2013.
- [48] D. De Nigris, L. Mercier, R. Del Maestro, D. Louis Collins, and T. Arbel. Hierarchical Multimodal Image Registration Based on Adaptive Local Mutual Information. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2010*, pages 643–651. Centre for Intelligent Machines, McGill University, Springer Berlin / Heidelberg, 2010.
- [49] F. Dellaert and R. Collins. Fast Image-Based Tracking by Selective Pixel Integration. Technical report, Corfu, Greece, Sept. 1999.
- [50] S. Dieterich, J. Rodgers, and R. Chan. Radiosurgery. *Image-Guided Interventions*, (Chapter 16):461–500, 2008.
- [51] C. Elliott, D. Collins, D. Arnold, and T. Arbel. Temporally consistent probabilistic detection of new multiple sclerosis lesions in brain mri. 2013.
- [52] A. C. Evans, W. Dai, D. L. Collins, P. Neelin, and S. Marrett. Warping of a computerized 3D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis. In *Proceedings of SPIE – Volume 1445 Medical Imaging V: Image Processing*, pages 236–246, 1991.
- [53] B. Fischer and J. Modersitzki. A unified approach to fast image registration and a new curvature based registration technique. *Linear Algebra and its applications*, 2004.
- [54] B. Fischer and J. Modersitzki. Ill-posed medicine—an introduction to image registration. *Inverse problems*, 2008.
- [55] J. Fitzpatrick and J. West. The distribution of target registration error in rigid-bodypoint-based registration. Jan. 2001.

- [56] V. Fonov, A. C. Evans, K. Botteron, C. R. Almlí, R. C. McKinstry, and D. L. Collins. Unbiased average age-appropriate atlases for pediatric studies. *54(1):313–327*, Jan. 2011.
- [57] M. Freiman, M. Werman, and L. Joskowicz. A curvelet-based patient-specific prior for accurate multi-modal brain image rigid registration. *Medical Image Analysis*, 15(1):125–132, 2011.
- [58] K. J. Friston, S. Williams, R. Howard, R. S. J. Frackowiak, and R. Turner. Movement-Related effects in fMRI time-series. *Magnetic Resonance in Medicine*, 35(3):346–355, 1996.
- [59] K. A. Ganser, H. Dickhaus, R. Metzner, and C. R. Wirtz. A deformable digital brain atlas system according to Talairach and Tournoux. *Medical Image Analysis*, 2004.
- [60] J. C. Gee, M. Reivich, and R. Bajcsy. Elastically Deforming 3D Atlas to Match Anatomical Brain Images. *Journal of Computer Assisted Tomography*, 17(2):225, 1993.
- [61] D. G. Gobbi and T. M. Peters. Interactive intra-operative 3D ultrasound reconstruction and visualization. *Medical Image Computing and Computer-Assisted ...*, 2002.
- [62] A. A. Goshtasby. 2-D and 3-D image registration: for medical, remote sensing, and industrial applications, 2005.
- [63] A. Gronningsaeter, A. Kleven, S. Ommedal, T. E. Aarseth, T. Lie, F. Lindseth, T. Langø, and G. Unsgård. SonoWand, an Ultrasound-based Neuronavigation System. *Neurosurgery*, 47(6):1373, Dec. 2000.
- [64] E. Haber and J. Modersitzki. Numerical methods for volume preserving image registration. *Inverse problems*, 20(5):1621–1638, Oct. 2004.
- [65] E. Haber and J. Modersitzki. Beyond Mutual Information: A Simple and Robust Alternative. *Bildverarbeitung für die Medizin 2005*, (Chapter 72):350–354, 2005.
- [66] J. V. Hajnal, N. Saeed, E. J. Soar, A. Oatridge, I. R. Young, and G. M. Bydder. A Registration and Interpolation Procedure for Subvoxel Matching of Serially Acquired MR Images. *Journal of Computer Assisted Tomography*, 19(2):289, 1995.
- [67] N. Hansen and A. Ostermeier. Completely Derandomized Self-Adaptation in Evolution Strategies. *Evol. Comput.*, 9:159–195, June 2001.

- [68] M. Heinrich, J. Schnabel, F. Gleeson, and M. Brady. Non-rigid multimodal medical image registration using optical flow and gradient orientation. *Medical Image Analysis and Understanding*, 2010.
- [69] M. P. Heinrich, M. Jenkinson, M. Bhushan, and T. Matin. Non-local shape descriptor: a new similarity metric for deformable multi-modal registration. ... *Image Computing and ...*, 2011.
- [70] P. Hellier, C. Barillot, E. Memin, and P. Perez. Hierarchical estimation of a dense deformation field for 3-D robust registration. *IEEE TMI*, 20(5):388–402, 2001.
- [71] A. O. Hero, B. Ma, O. J. J. Michel, and J. Gorman. Applications of entropic spanning graphs. *Signal Processing Magazine, IEEE*, 19(5):85–95, Sept. 2002.
- [72] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes. Medical image registration. *Physics in Medicine and Biology*, 46:R1–R45, 2001.
- [73] P. W. Hsu, R. W. Prager, A. H. Gee, and G. M. Treece. Freehand 3D ultrasound calibration: a review, 2009.
- [74] L. Ibanez, W. Schroeder, L. Ng, and J. Cates. *The ITK Software Guide: ITK V2.0*. Kitware Inc, Clifton Park, NY, USA, 2005.
- [75] I. Isgum, M. Staring, A. Rutten, M. Prokop, M. A. Viergever, and B. van Ginneken. Multi-Atlas-Based Segmentation With Local Decision Fusion Application to Cardiac and Aortic Segmentation in CT Scans. *IEEE TMI*, 28(7):1000–1010, July 2009.
- [76] D. Jaffray, P. Kupelian, T. Djemil, and R. M. Macklis. Review of image-guided radiation therapy. *dx.doi.org*, 7(1):89–103, Dec. 2006.
- [77] A. S. Jakola, K. S. Myrmel, R. Kloster, and S. H. Torp. Comparison of a strategy favoring early surgical resection vs a strategy favoring watchful waiting in low-grade gliomas. *Jama*, 2012.
- [78] S. Ji, Z. Wu, A. Hartov, D. W. Roberts, and K. D. Paulsen. Mutual-information-based image to patient re-registration using intraoperative ultrasound in image-guided neurosurgery. *Medical Physics*, 35(10):4612–4624, Oct. 2008.
- [79] S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23, 2004.
- [80] B. Kara cali. Information Theoretic Deformable Registration Using Local Image Information. *IJCV*, 72:219–237, 2007.

- [81] Z. Karimaghloo, H. Rivaz, D. L. Arnold, D. L. Collins, and T. Arbel. Adaptive Voxel, Texture and Temporal Conditional Random Fields for Detection of Gad-Enhancing Multiple Sclerosis Lesions in Brain MRI. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, pages 543–550. Springer, 2013.
- [82] Y. Keller and A. Averbuch. Fast motion estimation using bi-directional gradient methods. *IEEE Transactions on Image Processing*, 13(8), 2004.
- [83] P. J. Kelly, J. S. Loeffler, M. R. Bussiere, and C. Yu. An analysis of the accuracy of the Cyberknife: A robotic nameless stereotactic radiosurgical system-Comments, 2003.
- [84] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, 46(3):786–802, 2009.
- [85] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim. elastix: a toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205, 2010.
- [86] R. K. S. Kwan, A. C. Evans, and G. B. Pike. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Transactions on Medical Imaging*, 18(11):1085–1097, 1999.
- [87] M. Letteboer, P. Willems, M. Viergever, and W. Niessen. Brain shift estimation in image-guided neurosurgery using 3-D ultrasound. *Biomedical Engineering, IEEE Transactions on*, 52(2):268–276, 2005.
- [88] M. M. J. Letteboer, P. Willems, M. A. Viergever, and W. J. Niessen. Brain Shift Estimation in Image-Guided Neurosurgery Using 3D Ultrasound. *IEEE Transactions on Biomedical Engineering*, 52(2), 2004.
- [89] S. Liao and A. C. S. Chung. Non-rigid Image Registration with Uniform Gradient Spherical Patterns. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009*, pages 696–704. Springer Berlin Heidelberg, Berlin, Heidelberg, Jan. 2009.
- [90] T. Lindeberg. *Scale-space theory in computer vision*. Springer, Dec. 1993.
- [91] F. Lindseth, T. Langø, J. Bang, and T. A. Nagelhus Hernes. Accuracy evaluation of a 3D ultrasound-based neuronavigation system. *Computer Aided Surgery*, 7(4):197–222, Jan. 2002.

- [92] F. Lindseth, T. Langø, T. Selbekk, R. Hansen, I. Reinertsen, C. Askeland, O. Solheim, G. Unsgård, R. Mårvik, and T. A. N. Hernes. Ultrasound-Based Guidance and Therapy. 2013.
- [93] F. Lindseth, G. A. Tangen, T. Langø, and J. Bang. Probe calibration for freehand 3-D ultrasound. *Ultrasound in medicine & biology*, 29(11):1607–1623, Nov. 2003.
- [94] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens. Nonrigid Image Registration Using Conditional Mutual Information. *IEEE TMI*, 29(1):19–29, 2010.
- [95] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, Nov. 2004.
- [96] B. D. Lucas and T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. In *Proceedings of Image Understanding Workshop*, pages 121–130, 1981.
- [97] F. Maes, D. Vandermeulen, and P. Suetens. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 3(4):373–386, 1999.
- [98] J. B. A. Maintz and M. A. Viergever. A Survey of medical Image registration. *Medical Image Analysis*, 2(1):1–36, 1998.
- [99] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. PET-CT image registration in the chest using free-form deformations. *IEEE TMI*, 22(1):120–128, Jan. 2003.
- [100] J. C. Mazziotta, A. W. Toga, A. Evans, P. Fox, and J. Lancaster. A probabilistic atlas of the human brain: theory and rationale for its development the international consortium for brain mapping (ICBM). *NeuroImage*, 1995.
- [101] M. Mellor and M. Brady. Non-rigid Multimodal Image Registration Using Local Phase. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2004*, pages 789–796. 2004.
- [102] L. Mercier. *Ultrasound-guided brain tumor resection*. PhD thesis, McGill University, Montreal.
- [103] L. Mercier, R. F. Del Maestro, K. Petrecca, D. Araujo, C. Haegelen, and D. L. Collins. Online database of clinical MR and ultrasound images of brain tumors. *Medical Physics*, 39(6):3253–3261, May 2012.

- [104] L. Mercier, R. F. Del Maestro, K. Petrecca, A. Kochanowska, S. Drouin, C. X. B. Yan, A. L. Janke, S. J.-S. Chen, and D. L. Collins. New prototype neuronavigation system based on preoperative imaging and intraoperative freehand ultrasound: system description and validation. *International Journal of Computer Assisted Radiology and Surgery*, 6(4):507–522, 2011.
- [105] L. Mercier, V. Fonov, C. Haegelen, R. F. Del Maestro, K. Petrecca, and D. L. Collins. Comparing two approaches to rigid registration of three-dimensional ultrasound and magnetic resonance images for neurosurgery. *International Journal of Computer Assisted Radiology and Surgery*, 7(1):125–136, Jan. 2012.
- [106] L. Mercier, T. Langø, F. Lindseth, and L. D. Collins. A review of calibration techniques for freehand 3-D ultrasound systems. *Ultrasound in medicine & biology*, 31(2):143–165, Feb. 2005.
- [107] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(10):1615–1630, Oct. 2005.
- [108] M. I. Miller, A. Trouvé, and L. Younes. On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, 4(1):375–405, Aug. 2002.
- [109] K. Murphy, B. van Ginneken, S. Klein, and M. Staring. Semi-automatic construction of reference standards for evaluation of image registration. *Medical image ...*, 2011.
- [110] K. Murphy, B. van Ginneken, J. M. Reinhardt, S. Kabus, K. Ding, X. Deng, and J. P. W. Pluim. Evaluation of Methods for Pulmonary Image Registration: The EMPIRE10 Study. In *Grand Challenges in Medical Image Analysis*, 2010.
- [111] A. Nabavi, N. Hata, D. Gering, E. Chatzidakis, M. Leventon, N. Weisenfeld, R. Perogolizzi, K. Oge, P. Black, F. Jolesz, and R. Kikinis. Image guided neurosurgery visualization of brain shift. pages 17–26, 1999.
- [112] R. N. Nagel and A. Rosenfeld. Ordered search techniques in template matching. *Proceedings of the IEEE*, 60(2):242–244, Feb. 1972.
- [113] F. P. M. Oliveira and J. M. R. S. Tavares. Medical image registration: a review. *Computer methods in biomechanics and biomedical engineering*, 17(2):73–93, 2014.
- [114] B. N. Oreshkin and T. Arbel. Uncertainty Driven Probabilistic Voxel Selection for Image Registration. *IEEE TMI*, 32(10):1777–1790, 2013.
- [115] R. P, S. DW, and A. JR. Image-guided radiosurgical ablation of intra- and extra-cranial lesions. *Technology in cancer research & treatment*, 5(4):421–428, Aug. 2006.

- [116] X. Pennec, P. Cachier, and N. Ayache. Understanding the "Demon's Algorithm": 3D Non-rigid Registration by Gradient Descent. In *MICCAI '99: Proceedings of the Second International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer-Verlag, Sept. 1999.
- [117] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes. A comparison of similarity measures for use in 2-D-3-D medical image registration. *Medical Imaging, IEEE Transactions on*, 17(4):586–595, 1998.
- [118] J. Pluim, J. Maintz, and M. A. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE TMI*, 22(8):1004–1016, July 2003.
- [119] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Image registration by maximization of combined mutual information and gradient information. *IEEE TMI*, 19(8):809–814, Aug. 2000.
- [120] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual-Information Based Registration of Medical Images: A Survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, 2003.
- [121] T. C. Poon and R. N. Rohling. Comparison of calibration methods for spatial tracking of a 3-D ultrasound probe. *Ultrasound in medicine & biology*, 2005.
- [122] K. Popuri, D. Cobzas, and M. Jägersand. A Variational Formulation for Discrete Registration. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, pages 187–194. Springer Berlin Heidelberg, 2013.
- [123] S. J. Reeves and R. Hezar. Selection of observations in magnetic resonance spectroscopic imaging. In *ICIP '95: Proceedings of the 1995 International Conference on Image Processing (Vol. 1)*. IEEE Computer Society, Oct. 1995.
- [124] I. Reinertsen, F. Lindseth, G. Unsgaard, and D. L. Collins. Clinical validation of vessel-based registration for correction of brain-shift. *Medical Image Analysis*, 2007.
- [125] M. H. T. Reinges, H. H. Nguyen, T. Krings, B. O. Hütter, V. Rohde, and J. M. Gilsbach. Course of brain shift during microsurgical resection of supratentorial cerebral lesions: limits of conventional neuronavigation. *Acta Neurochirurgica*, 146(4):369–377, Apr. 2004.
- [126] M. Reuter, N. J. Schmansky, H. D. Rosas, and B. Fischl. Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, 61(4):1402–1418, July 2012.

- [127] H. Rivaz, Z. Karimaghloo, and D. L. Collins. Self-similarity weighted mutual information: A new nonrigid image registration metric. *Medical Image Analysis*, 18(2):343–358, Feb. 2014.
- [128] H. Rivaz, Z. Karimaghloo, V. S. Fonov, and D. L. Collins. Nonrigid Registration of Ultrasound and MRI Using Contextual Conditioned Mutual Information. *Medical Imaging, IEEE Transactions on*, 33(3):708–725, 2014.
- [129] D. W. Roberts, A. Hartov, F. E. Kennedy, M. I. Miga, and K. D. Paulsen. Intraoperative Brain Shift and Deformation: A Quantitative Analysis of Cortical Displacement in 28 Cases. *Neurosurgery*, 43(4):749, Oct. 1998.
- [130] A. Roche, G. Malandain, N. Ayache, and S. Prima. Towards a better comprehension of similarity measures used in medical image registration. In *Proceedings of the 2nd International conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'99)*, Cambridge, UK, 1999. Springer.
- [131] A. Roche, X. Pennec, G. Malandain, and N. Ayache. Rigid Registration of 3D Ultrasound with MR Images: a New Approach Combining Intensity and Gradient Information. *IEEE TMI*, 20:1038–1049, 2001.
- [132] T. Rohlfing. Image Similarity and Tissue Overlaps as Surrogates for Image Registration Accuracy: Widely Used but Unreliable. *IEEE TMI*, 31(2):153–163, 2012.
- [133] T. Rohlfing, C. R. Maurer Jr, D. A. Bluemke, and M. A. Jacobs. Volume-Preserving Nonrigid Registration of MR Breast Images Using Free-Form Deformation With an Incompressibility Constraint. *IEEE Transactions on Medical Imaging*, 22(6):730–741, 2003.
- [134] T. Rohlfing, E. V. Sullivan, and A. Pfefferbaum. Subject-matched templates for spatial normalization. *Medical Image Computing and ...*, 2009.
- [135] R. Rohling, A. Gee, and L. Berman. A comparison of freehand three-dimensional ultrasound reconstruction techniques. *Medical Image Analysis*, 1999.
- [136] R. N. Rohling. 3D freehand ultrasound: reconstruction and spatial compounding. *Th se de Doctorat*, 1998.
- [137] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
- [138] D. Russakoff, C. Tomasi, T. Rohlfing, and Calvin Jr. Image Similarity Using Mutual Information of Regions. In *Computer Vision - ECCV 2004*, pages 596–607, 2004.

- [139] M. R. Sabuncu and P. J. Ramadge. *Gradient based nonuniform subsampling for information-theoretic alignment methods*, volume 1. IEEE, 2004.
- [140] M. R. Sabuncu, B. T. T. Yeo, K. Van Leemput, B. Fischl, and P. Golland. A Generative Model for Image Segmentation Based on Label Fusion. *IEEE TMI*, 29(10):1714–1729, Oct. 2010.
- [141] M. R. Sabuncu, B. T. T. Yeo, K. Van Leemput, T. Vercauteren, and P. Golland. Asymmetric Image-Template Registration. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009*, pages 565–573. Springer Berlin Heidelberg, Berlin, Heidelberg, Jan. 2009.
- [142] C. A. Saether, M. Torsteinen, S. H. Torp, S. Sundstrøm, G. Unsgård, and O. Solheim. Did Survival Improve after the Implementation of Intraoperative Neuronavigation and 3D Ultrasound in Glioblastoma Surgery? A Retrospective Analysis of 192 Primary Operations. *Journal of Neurological Surgery Part A: Central European Neurosurgery*, 73(02):073–078, Mar. 2012.
- [143] N. Sanai and M. S. Berger. Operative techniques for gliomas and the value of extent of resection. *Neurotherapeutics*, 6(3):478–486, July 2009.
- [144] J. A. Schnabel, C. Tanner, A. D. Castellano-Smith, A. Degenhard, M. O. Leach, D. R. Hose, D. L. G. Hill, and D. J. Hawkes. Validation of nonrigid image registration using finite-element methods: application to breast MR images. *IEEE TMI*, 22(2):238–247, Feb. 2003.
- [145] R. Shams, R. A. Kennedy, P. Sadeghi, and R. Hartley. Gradient Intensity-Based Registration of Multi-Modal Images of the Brain. *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8, 2007.
- [146] R. Shams, P. Sadeghi, and R. A. Kennedy. Gradient Intensity: A New Mutual Information-Based Registration Method. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1–8. IEEE, 2007.
- [147] D. W. Shattuck, M. Mirza, V. Adisetiyo, and C. Hojatkashani. Construction of a 3D probabilistic atlas of human cortical structures. *NeuroImage*, 2008.
- [148] E. Shechtman and M. Irani. Matching Local Self-Similarities across Images and Videos. *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.
- [149] R. Sibson. Studies in the Robustness of Multidimensional Scaling: Procrustes Statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 40(2):pp. 234–238, 1978.

- [150] P. J. Slomka and R. P. Baum. Multimodality image registration with software: state-of-the-art. *European journal of nuclear medicine and ...*, 2009.
- [151] O. V. Solberg, F. Lindseth, H. Torp, R. E. Blake, and T. A. Nagelhus Hernes. Freehand 3D Ultrasound Reconstruction Algorithms—A Review. *Ultrasound in medicine & biology*, 33(7):991–1009, July 2007.
- [152] O. Solheim, T. Selbekk, A. S. Jakola, and G. Unsgård. Ultrasound-guided operations in unselected high-grade gliomas—overall results, impact of image quality and patient selection. *Acta Neurochirurgica*, 2010.
- [153] M. P. Sormani, L. Bonzano, L. Roccatagliata, G. R. Cutter, G. L. Mancardi, and P. Bruzzi. Magnetic resonance imaging as a potential surrogate for relapses in multiple sclerosis: A meta-analytic approach. *Annals of Neurology*, 65(3):268–275, Mar. 2009.
- [154] R. Sprengel, K. Rohr, and H. S. Steidl. Thin-plate spline approximation for image registration. In *Proceedings of the 18th International Conference of the IEEE Engineering in Medicine and Biology Society*, 1996.
- [155] M. Staring, S. Klein, and J. P. W. Pluim. A rigidity penalty term for nonrigid registration. *Medical Physics*, 34:4098–+, 2007.
- [156] M. Staring, U. A. van der Heide, S. Klein, M. A. Viergever, and J. P. W. Pluim. Registration of Cervical MRI Using Multifeature Mutual Information. *IEEE TMI*, 28(9):1412–1421, 2009.
- [157] L. H. Stieglitz, J. Fichtner, R. Andres, P. Schucht, A.-K. Krähenbühl, A. Raabe, and J. Beck. The Silent Loss of Neuronavigation Accuracy: A Systematic Retrospective Analysis of Factors Influencing the Mismatch of Frameless Stereotactic Systems in Cranial Neurosurgery. *Neurosurgery*, 72(5):796–807, May 2013.
- [158] N. K. Subbanna, D. Precup, D. L. Collins, and T. Arbel. Hierarchical Probabilistic Gabor and MRF Segmentation of Brain Tumours in MRI Volumes. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, pages 751–758. Springer, 2013.
- [159] R. Szeliski. Image alignment and stitching: A tutorial. Technical report, Dec. 2004.
- [160] P. Thevenaz and M. Unser. An efficient mutual information optimizer for multiresolution image registration. In *Proceedings of the International Conference on Image Processing (ICIP 98)*, 1998.
- [161] P. Thevenaz and M. Unser. Optimization of mutual information for multiresolution image registration. *Image Processing, IEEE Transactions on*, 9(12):2083–2099, 2000.

- [162] J. P. Thirion. Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical Image Analysis*, 2(3):243–260, Sept. 1998.
- [163] A. Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–221, 1998.
- [164] M. Unser, A. Aldroubi, and M. Eden. B-Spline Signal Processing: Part I—Theory. *IEEE Transactions on Signal Processing*, 41(2):821–833, 1993.
- [165] G. Unsgaard, S. Ommedal, T. Muller, A. Gronningsaeter, and T. A. Nagelhus Hernes. Neuronavigation by Intraoperative Three-dimensional Ultrasound: Initial Experience during Brain Tumor Resection. *Neurosurgery*, 50(4):804, Apr. 2002.
- [166] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. Non-parametric Diffeomorphic Image Registration with the Demons Algorithm. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, 4792(Chapter 39):319–326, 2007.
- [167] P. Viola and W. M. Wells III. Alignment by Maximization of Mutual Information. In *Proceedings of the 5th IEEE International Conference on Computer Vision (ICCV1995)*, pages 16–23, Cambridge, MA, USA, June 1995. IEEE Computer Society.
- [168] G. Wahba. Spline models for observational data, 1990.
- [169] H. Wang, L. Dong, J. O'Daniel, R. Mohan, A. S. Garden, K. K. Ang, D. A. Kuban, M. Bonnen, J. Y. Chang, and R. Cheung. Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy. *Physics in Medicine and Biology*, 50(12):2887–2905, June 2005.
- [170] H. Wang, J. W. Suh, S. R. Das, J. B. Pluta, C. Craige, and P. A. Yushkevich. Multi-Atlas Segmentation with Joint Label Fusion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(3):611–623, 2013.
- [171] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35–51, Mar. 1996.
- [172] J. B. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr, R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. M. F. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, P. F. Hemler, S. Napel, T. S. Sumanaweera, B. A. Harkness, D. L. Hill, C. Studholme, G. Malandain, X. Pennec, M. E. Noz, G. Q. Maguire, Jr, M. Pollack, C. A. Pelizzari, R. A. Robb, D. P. Hanson, and R. P. Woods. Comparison and evaluation of retrospective intermodality image registration techniques. 2710(1):332–347, 1996.

- [173] C. Wu, P. E. Murtha, and B. Jaramaz. Femur statistical atlas construction based on two-level 3D non-rigid registration. *Computer Aided Surgery*, 2009.
- [174] P. A. Yushkevich, B. B. Avants, S. R. Das, J. Pluta, M. Altinay, and C. Craige. Bias in estimation of hippocampal atrophy using deformation-based morphometry arises from asymmetric global normalization: An illustration in ADNI 3 T MRI data. *NeuroImage*, 50(2):434–445, Apr. 2010.
- [175] P. A. Yushkevich, B. B. Avants, J. Pluta, S. Das, and D. Minkoff. A high-resolution computational atlas of the human hippocampus from postmortem magnetic resonance imaging at 9.4 T. *NeuroImage*, 2009.
- [176] B. Zitova and J. Flusser. Image Registration Methods: A Survey. *Image and Vision Computing*, 21(11):977–1000, 2003.