Advances in Groupwise Image Registration

Ontwikkelingen in groepsgewijze beeldregistratie

Thesis

to obtain the degree of Doctor of Engineering Sciences (Doctor in de ingenieurswetenschappen) at Vrije Universiteit Brussel and the degree of Doctor at Erasmus Universiteit Rotterdam

by command of the rector magnificus Prof. dr. A.L. Bredenoord, Erasmus University Rotterdam, and in accordance with the decision of both Doctorate Boards.

The public defence shall be held on Wednesday May, 25th 2022 at 10:30 at Rotterdam.

by

Mathias Luc Julia Polfliet
born in Vilvoorde, Belgium

Erasmus University Rotterdam

VRIJE UNIVERSITEIT BRUSSEL

Doctoral committee

Promotors:

Prof. Dr. Ir. J. Vandemeulebroucke
Prof. Dr. W.J. Niessen

Other members:

Prof. Dr. Ir. R. Pintelon
Prof. Dr. A. Dooms
Prof. Dr. M. de Bruijne
Prof. Dr. D. Rueckert
Prof. Dr. F. Maes

Co-promotor:

Dr. Ir. S. Klein

I am the master of my fate,
I am the captain of my soul.

Contents

Abstract

This thesis deals with advances in groupwise image registration. Image registration remains an important task in medical image analysis. Whereas most methods are designed for the registration of two images (pairwise registration), there is an increasing interest in simultaneously aligning more than two images using groupwise registration given the increasing availability of medical imaging data, both at the individual and the population level. Groupwise image registration has shown promise in a number of applications dealing with large quantities of data, among others to increase registration accuracy and robustness, to improve the transformation smoothness and to reduce the methodological bias compared to pairwise registrations. However, directly comparing groupwise registrations to conventional repeated pairwise registrations is difficult due to several confounding factors impacting the algorithm. In this thesis, as a first contribution, we rigorously evaluate two registration methodologies in several experiments and investigate the differences in performance. Secondly, we fill a gap in current literature on efficient (dis)similarity measures for multimodal groupwise image registration. These two contributions are distributed over four chapters.

In **Chapter 3**, we investigate several registration approaches for the alignment of CT and MRI acquisitions of the mandible in patients with oral squamous cell carcinoma. A comparison is made between rigid and non-rigid approaches with symmetric and asymmetric transformation strategies. The results suggest improved performance in terms of registration accuracy for a symmetric transformation strategy compared to an asymmetric approach, however, the differ-

ences were not statistically significant (p=0.054). For this clinical application, we conclude that a rigid registration method is the recommended approach.

In **Chapter 4**, an investigation is performed on different template images for groupwise registrations based on mutual information. Here, template images are employed as a representative image to compare every image in the group to (in terms of its (dis)similarity). We show that the entropy of the template image can have a counter-intuitive contribution to the global dissimilarity value. Additionally, we show that equivalent performance in terms of registration accuracy can be achieved between groupwise and repeated pairwise approaches.

In **Chapter 5**, a novel similarity measure is introduced for multimodal groupwise registration. The conditional template entropy measures the negated average of the pairwise conditional entropy of each image of the group and a template image, which is constructed based on principal component analysis. We show improved or equivalent performance in terms of accuracy compared to other state-of-the-art (dis)similarity measures for multimodal groupwise registration and repeated pairwise registration. Furthermore, groupwise registration vastly outperform repeated pairwise registration in terms of transitive error, a measure which can be interpreted as a measure for the consistency of the transformations in a groupwise setting.

In **Chapter 6**, to further improve on the efficiency of multimodal groupwise registration, we propose a novel dissimilarity measure which is especially adept at registering large groups of images. The dissimilarity measure is formulated as the second smallest eigenvalue of the generalized eigenvalue problem posed in the description of Laplacian eigenmaps. We show little dependence of the measure in terms of computation time with respect to the number of images in the group, and equivalent or improved performance in terms of registration accuracy compared to state-of-the-art groupwise (dis)similarity measures.

To summarize, in this work we evaluate groupwise approaches compared to repeated pairwise approaches and show mostly equivalent performance in terms of registration accuracy and robustness and an improved transitivity for groupwise registration. Furthermore, we recommend to use the proposed dissimilarity measure based on Laplacian eigenmaps for large groups of images given its superior or equivalent registration accuracy compared to other measures but superior scaling in terms of execution time with respect to the number of images in the group.

Samenvatting

Deze thesis behandelt ontwikkelingen in groepsgewijze beeldregistratie. Beeldregistratie blijft een relevante en belangrijke taak bij medische beeldanalyse. Daar waar de meeste methoden ontwikkelt zijn voor de registratie van twee beelden (paarsgewijze registratie), is er een groeiende interesse in de gelijktijdige alignatie van meer dan twee beelden. Dit omwille van de groeiende beschikbaarheid van medische beelden, zowel op een individueel niveau als op het niveau van een bevolking. Groepsgewijze registratie behaalde reeds beloftevolle resultaten in applicaties met grote hoeveelheden aan data, onder andere om de registratie nauwkeurigheid en robuustheid te verhogen, de plausibiliteit van de resulterende transformatie te verhogen en om methodologische bias te verminderen vergeleken met paarsgewijze registraties. Echter, de directe vergelijking maken tussen paarsgewijze en groepsgewijze registraties is moeilijk omwille van verscheidene confounding variabelen die de algoritmes beïnvloeden. In deze thesis, als een eerste contributie, vergelijken we de twee registratie methodologieën rigoureus met elkaar in verscheidene experimenten en onderzoeken we hun verschillen. Als tweede contributie onderzoeken we efficiënte metrieken voor multimodale groepsgewijze registratie. Deze twee contributies zijn verdeeld over vier hoofdstukken.

In **Hoofdstuk 3** onderzoeken we verschillende registratiemethodes voor de alignatie van CT en MRI acquisities van de mandibula voor patiënten met oraal plaveiselcelcarcinoom. Een vergelijking is opgesteld tussen rigide en niet-rigide methoden met een symmetrische en asymmetrische transformatiestrategie. De resultaten suggereren een verbetering in termen van registratienauwkeurigheid

voor de symmetrische strategie, echter de verschillen zijn niet statistisch significant. Voor deze klinische applicatie, concluderen we dat een rigide registratiemethode de aanbevolen manier is.

In **Hoofdstuk 4** voeren we een onderzoek uit naar verschillende template-beelden voor groepsgewijze registratie op basis van mutuele informatie. Hier worden template-beelden gebruikt als een representatief beeld om elk beeld in de groep mee te vergelijken (in termen van de registratiemetriek). we tonen dat de entropie van het template-beeld een contra-intuitieve bijdrage kan leveren bij de totale metriekwaarde. Verder tonen we dat gelijkaardige resultaten behaald worden in termen van registratienauwkeurigheid wanneer we groepsgewijze en herhalende paarsgewijze methoden vergelijken.

In **Hoofdstuk 5** introduceren we een nieuw metriek voor multimodale groepsgewijze registratie. De conditionele template entropie meet het negatieve gemiddelde van de paarsgewijze conditionele entropie van elk beeld in de groep met een template beeld, geconstrueerd met behulp van principale component analyse. We tonen aan dat deze nieuwe metriek verbeterde of gelijkaardige registratienauwkeurigheid oplevert in vergelijking met andere metrieken voor multimodale groepsgewijze registratie en herhalende paarsgewijze registratie. In termen van transitiviteit, een maat die kan geïnterpreteerd worden als een maat voor de consistentie van de transformaties voor een groep van beelden, tonen we dat groepsgewijze registraties een meerwaarde hebben vergeleken met herhalende paarsgewijze registraties.

In **Hoofdstuk 6** gaan we verder met het verbeteren van de efficiëntie van multimodale groepsgewijze registraties en stellen we een nieuwe metriek voor die toegespitst is op de registratie van grote groepen van beelden. De metriek is opgesteld als de tweede kleinste eigenwaarde van het gegeneralizeerd eigenwaardenprobleem geponeerd binnen de context van Laplaciaanse eigenmappen. We tonen dat de benodigde rekentijd voor de voorgestelde metriek amper afhankelijk is van de hoeveelheid beelden die geregistreerd moeten worden. Verder tonen we gelijkaardige of verbeterde registratienauwkeurigheid in vergelijking met andere groepsgewijze metrieken.

Ter conclusie, in dit werk evalueren we groepsgewijze registratiemethoden ten opzichte van herhalende paarsgewijze registraties en we demonstreren voornamelijk gelijkaardige resultaten in termen van registratienauwkeurigheid en robuustheid en verbeterde resultaten in termen van transiviteit voor groepsgewijze registraties. Verder ontwerpen we een nieuwe metriek voor groepsgewijze multimodale registratie gebaseerd op Laplaciaanse eigenmappen en raden we zijn verder gebruik aan. Dit omwille van zijn verbeterde of gelijkaardige resultaten in termen van registratienauwkeurigheid in vergelijking met andere

metrieken, maar superieure schaling van zijn computationele rekentijd in functie van het aantal beelden in de groep.

CHAPTER 1

Introduction

Image registration is the process of finding the set of transformations which most optimally aligns a corresponding set of images (see Fig. 1.1). It is typically formulated as an optimization problem where an objective function is minimized or maximized with respect to the sought-after transformations.

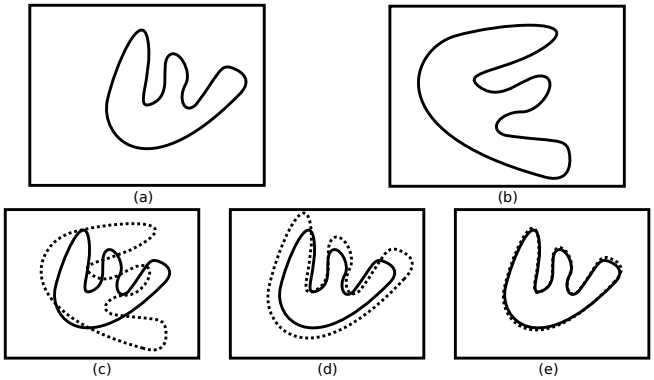
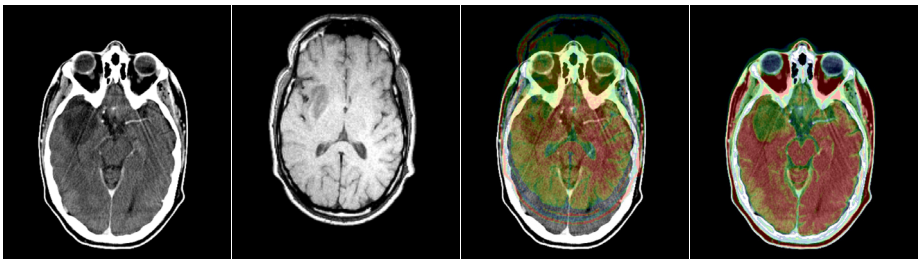


Figure 1.1: Top - (a) and (b) are two images of the same object which are misaligned and require registration. Bottom - the two images overlayed, (c) after an initial translation, (d) followed by a rotation, (e) and a scaling.

1.1 Biomedical image registration

Medical images stand in sharp contrast to images we come across in our everyday life. Medical imaging deals primarily with volumetric acquisitions where the voxel size is quantified in physical coordinates. Where the former property increases the complexity of algorithms due to scalability issues, the latter property allows for a better initialization of the alignment problem and reduces the complexity of the registration task by reducing the degrees of freedom. Additionally, the imaging modalities available in the medical field differ vastly from those available in other research fields, where the optical spectrum is typically the main focus. Medical imaging include technologies using electromagnetic waves, spanning a large range of frequencies (from long wavelengths in magnetic resonance imaging to short wavelengths in computed tomography) and acoustic waves (in ultrasound). A complete overview of the medical imaging modalities would lead us too far and the reader is referred to Bushberg and Boone (2011); Cobbold (2006) and Hsieh et al. (2009) for a thorough review.



(a) (b) (c) (d)

Figure 1.2: Medical images of the brain in a subject with a malignant lesion. (a) A computed tomography (CT) acquisition. (b) A magnetic resonance (MR) acquisition. (c) A fusion image overlaying the CT and MR acquisition prior to registration. (d) A fusion image overlaying the CT and MR acquisition after registration is performed.

When applied to medical imaging, image registration is coined *biomedical* image registration. By aligning two images to the same frame of reference, it allows physicians to quantitatively compare the image intensities at the same anatomical location in two different medical images, for example to measure tumor growth, lung ventilation or bone porosity. Those images could differ in terms of subject, acquisition date, acquisition time and/or imaging modality. Since the subject is not positioned in exactly the same manner in these two images and the object of interest may differ in size and shape, such a quantitative comparison can be difficult without prior alignment (see Fig. 1.2). In addition to guiding or aiding physicians in making better decisions, image registration might also be needed as a preprocessing step in an (semi-) automatic image analysis pipeline that requires an image to be aligned to a template or atlas image. We provide the following overview of applications of biomedical image registration, which is undoubtedly incomplete given the vast number of works.

Image registration can be used to perform follow-up analyses or comparisons for disease progression or treatment assessment (Brock et al., 2006; Charlton et al., 2010; Giesel et al., 2009; Giles et al., 2014; Gorbunova et al., 2008; Lynch et al., 2001; Nishiyama et al., 2015; Staring et al., 2007; Van Assche et al., 2007). In treatment planning, numerous applications can be found for radiotherapy planning (Foskey et al., 2005; Ireland et al., 2007; Leibfarth et al., 2013; Lu et al., 2006; Mackie et al., 2003; Nishioka et al., 2002; König et al., 2016; Oh and Kim, 2017) or image guided surgery (Gerber et al., 2014; Huang et al.,

2005; Pennec et al., 2003; Risholm et al., 2011). Images from which quantitative parameter maps, such as diffusion or perfusion maps, can be extracted, could require or benefit from prior image registration (Galbán et al., 2012; Hallack et al., 2014; Guyader et al., 2015; Ng et al., 2011). In computational anatomy, where the shape and geometry of anatomy is studied, image registration plays a key role to build the manifolds on which the anatomies reside (Adaszewski et al., 2013; Joshi et al., 2004; Ma et al., 2008; Miller et al., 2009). Image registration can also be employed to perform atlas-based segmentation, where an intensity atlas image is registered to a target image of interest onto which the corresponding annotated labeled atlas image is subsequently mapped (Aljabar et al., 2009; Bustamante et al., 2015; Cabezas et al., 2011; Isgum et al., 2009; Makropoulos et al., 2017; Michopoulou et al., 2009). Additionally, fusing or combining multiple images into a new image with improved quality (James and Dasarathy, 2014; Nemec et al., 2010; Wang and Ma, 2008) and image mosaicing or stitching (Ceranka et al., 2018; Wachinger et al., 2007, 2008) is often performed with image registration as well. Clearly, biomedical image registration remains an active area of research (see Fig. 1.3).

Reviews or surveys on medical image registration could provide the reader with a more complete overview of the works (Pluim et al., 2003; Zitova and Flusser, 2003; Sotiras et al., 2013; Oliveira and Tavares, 2014).

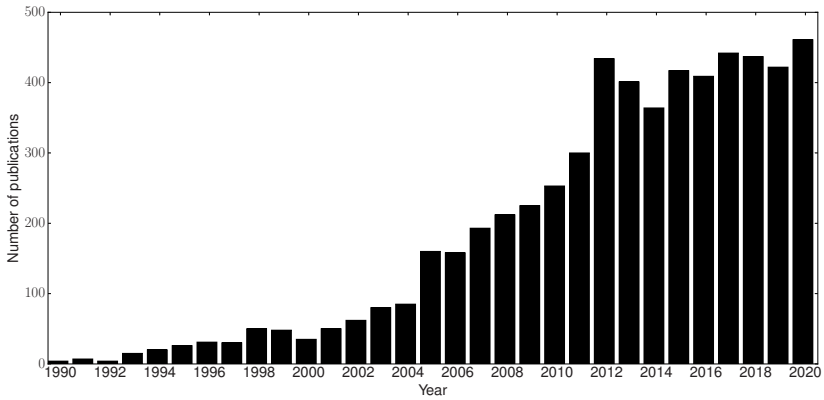


Figure 1.3: Yearly number of publications with “image registration” in the title or abstract in the Pubmed database, sorted by year.

1.2 Image registration in other fields

Image registration also holds value in other areas of research dealing with images. The following overview illustrates its widespread use.

In Krish et al. (2015), image registration is used to match partial to full fingerprints to perform automated fingerprint identification in the field of forensic science. Brown et al. (2008) imaged and digitally reassembled wall painting segments with image registration in the field of archaeology. In palaeography, the study of ancient and historical handwriting, image registration was employed to construct a deformable model which is subsequently used to date birch bark manuscripts (Sidorov, 2018). Furthermore, in augmented reality, image registration can be used to automate construction progress monitoring (Golparvar-Fard et al., 2009). Image registration has numerous applications in computer vision, it can be used to improve the quality and resolution of images (Irani and Peleg, 1991), to generate active shape or appearance models (Cootes et al., 1995, 2001) or to perform panoramic image stitching or mosaicing (Szeliski et al., 2007). Also in remote sensing a large number of applications exist, such as the quantification of land cover change (Dewan and Yamaguchi, 2009; Shalaby and Tateishi, 2007), to locate landslide locations (Cheng et al., 2004) or to perform image fusion (Simone et al., 2002).

1.3 Objectives of this thesis

Biomedical image registration, more often than not, deals with *two* images, meaning one image being registered to another image and this process is typically referred to as *pairwise* image registration. When more than two images are available the pairwise procedure is typically repeated. Here, one image is considered as a privileged image and all other images are registered to this privileged image. However, it might be beneficial to register all of them simultaneously, a process typically referred to as *groupwise* registration. This approach is the main subject of this work. Such an approach has been shown to carry both qualitative and quantitative advantages compared to a repeated pairwise approach. Qualitatively, a groupwise approach works generically for any number of images, does not require experiments to determine the 'best' privileged image and avoids the associated bias with such choice. Quantitatively, a groupwise approach can lead to improved accuracy, reliability or robustness (Cootes et al., 2004; Wachinger et al., 2007; Vandemeulebroucke et al., 2011; Wu et al., 2011; Yigitsoy et al., 2011; Huizinga et al., 2016; Royuela-del Val et al., 2016), increased smooth-

ness (Metz et al., 2011; Huizinga et al., 2016), reduced transitivity error (Geng et al., 2009; Metz et al., 2011; Polfliet et al., 2018) and enhanced downstream analysis in the image processing pipeline (Hamrouni et al., 2011; Huizinga et al., 2016; Sanz-Estébanez et al., 2017).

This work deals with advances in groupwise image registration and aims to accomplish two objectives. First, we will compare groupwise registrations to repeated pairwise registration in several experiments and evaluate their differences in terms of registration accuracy and transitivity. Secondly, we will propose novel methodologies to tackle multimodal groupwise registration efficiently, i.e. groupwise image registrations where not all included images originate from the same modality. This is achieved by applying techniques from the field of dimensionality reduction.

In the following chapter we will provide a methodological basis of biomedical image registration upon which we can situate our contributions and findings.

CHAPTER 2

Biomedical image registration

2.1 Formal definition

We start by formally defining image registration as an optimization problem:

$$\left(\hat{\mathcal{U}}_1, \ldots, \hat{\mathcal{U}}_n\right) = \arg \min_{(\mathcal{U}_1, \ldots, \mathcal{U}_n)} \mathcal{C}\left(R_1, \ldots, R_n; \mathcal{U}_1, \ldots, \mathcal{U}_n\right) \quad , \tag{2.1}$$

where $\mathcal{C}$ is the cost function which evaluates the misalignment of the image representations, R_i, given the displacement fields, $\mathcal{U}_i$. R_i is the employed representation for the underlying discrete image and $\mathcal{U}_i : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^d$ is the displacement field that maps the physical coordinates from the common d-dimensional reference domain, Ω, to a vector-valued displacement.

Note that recently some works formulate image registration as a function approximation task based on a learning algorithm (Balakrishnan et al., 2019; Sokooti et al., 2017). Obviously, in such approaches Eq. 2.1 does not hold. In this work, we will not deal with such approaches and revisit them later in the discussion.

2.2 Image registration approaches

Biomedical image registration algorithms can be categorized based on a number of aspects or criteria of the algorithm or the images it is trying to register. In this section we will provide an overview of possible categorizations. Note that the overview is not complete and categorizations are also possible based on properties other than those listed here. Furthermore, some registration schemes apply a combination of different properties and as such could be labeled as hybrid approaches.

The most important categorization that we will make is based on the number of images that require alignment as it is the main focus of this work. It will be tackled in section 2.3

2.2.1 Parametric and non-parametric transformations

Parametric and non-parametric registration approaches refer to the parameterization of the sought-after displacement field. Non-parametric approaches attempt to estimate the deformation or displacement field, $\mathcal{U}_i$, in Eq. (2.1) in each voxel or pixel directly. Some of the best known non-parametric approaches include the demons algorithm and its diffeomorphic equivalent (Thirion, 1998; Vercauteren et al., 2009).

Parametric approaches, however, derive the displacement field from a transformation function, $\mathcal{T}_{\boldsymbol{\mu}}$, parameterized by $\boldsymbol{\mu}$. As such, instead of optimizing the displacement field directly, an optimization is performed over the transformation parameters, $\boldsymbol{\mu}$. We can rewrite Eq. (2.1) to reflect this parameterization as

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{C}\left(R_1, \ldots, R_n; \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, \mathcal{T}_{\boldsymbol{\mu}_n}\right) \quad . \tag{2.2}$$

Herein is $\mathcal{T}_{\boldsymbol{\mu}_i} : \Omega \subset \mathbb{R}^d \rightarrow \Omega_i \subset \mathbb{R}^d$ the transformation that maps the coordinates from the common reference domain, Ω, to the domain of the ith image, Ω_i, and $\boldsymbol{\mu}_i$ are the parameters that define $\mathcal{T}_{\boldsymbol{\mu}_i}$. For the optimization, $\boldsymbol{\mu}$ is defined as a vector formed by the concatenation of all separate transformation parameters, $\boldsymbol{\mu}_i$.

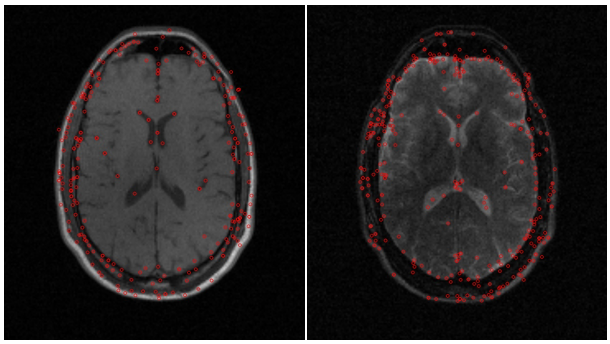
This work, specifically, only deals with parametric transformation models. As such, we will employ Eq. (2.2) to describe the remaining registration problems. All deformable transformations employed in this work are based on B-splines (Rueckert et al., 1999). They were chosen for their compact support and the associated computational advantages. It should be noted that B-splines produce smooth and continuous deformations which might not be suitable in certain applications where sliding motion is present (thoracic motion) or where certain structures vanish (removal of a organ). Alternatives for B-splines as a parametric deformable transformations model exist (Bookstein, 1989; Ristic and Brujic, 1997), but were not investigated.

2.2.2 Intensity-based and feature-based registration

Registration methods can be subdivided based on the representation that is employed for the images, R_i. Feature-based methods typically require two steps. First, the feature maps are calculated for the discrete images, which results in an abstraction for each image as a set of feature points (see Fig. 2.1), where the employed representation per image, R_i, becomes a set of d-dimensional points, f_i. Note that it is also possible to extract feature lines or areas from an image. Second, these sets of feature points are matched to each other minimizing some suitable cost function.

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{C}\left(f_1, \ldots, f_n; \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, \mathcal{T}_{\boldsymbol{\mu}_n}\right) \quad . \tag{2.3}$$

The main advantage of such feature-based methods is the computation time which is significantly reduced compared to intensity-based methods. Often used feature extractors are SIFT or SURF (Lowe, 2004; Bay et al., 2008).



(a) (b)

Figure 2.1: T1 and T2 weighted MRI images of the brain for which the SIFT points are extracted and highlighted in red. Images taken from Cocosco et al. (1997)

In intensity-based registration, an image is represented as an intensity function defined over a continuous domain. This intensity function can be obtained by interpolating the discrete image. Here, the image representation, R_i, becomes a function, $I_i : \Omega_i \subset \mathbb{R}^d \to \mathbb{R}$:

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{C}\left(I_1, \ldots, I_n; \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, \mathcal{T}_{\boldsymbol{\mu}_n}\right) \quad . \tag{2.4}$$

No abstraction of the image is constructed and all image information remains available during the registration process.

In this work, only intensity-based registration is tackled and we will continue to use Eq. (2.4) hereafter.

2.2.3 Stochastic and deterministic optimization

As image registration is typically defined as an optimization problem (the "arg min" part in Eq. 2.4), we could also categorize different registration schemes based on their optimization strategy. In a deterministic optimization scheme the resulting transformation is fully determined by the initialization of the registration. More specifically for intensity-based registration, a fixed number of samples are taken from the intensity function representing the images at, typically, predetermined coordinates to calculate the (dis)similarity measures. As

a result, the registration result is reproducible but might be biased due to the choice of initialization.

In stochastic optimization some form of randomness is introduced to the optimization. Such optimization is reported to be more robust to local optima and require less regularization (Klein et al., 2007; Sun et al., 2017). In intensity-based image registration this randomness is often achieved by randomly sampling a sparse subset of the image domain. Specifically in low-contrast regions of the images, the use of state-of-the-art deterministic approaches that assume the entire image domain is sampled, such as quasi-Newton and nonlinear conjugate gradient methods, may lead to unrealistic deformation fields due to the aggressive deformation field updates applied in each iteration. However, such issues could be alleviated with the use of a regularization term in the cost function optimization.

In this work, we only deal with stochastic optimization.

2.2.4 Monomodal and multimodal images

Biomedical images come in a vast range of modalities such as x-ray imaging, (cone-beam) computed tomography, ultrasound, positron emitted tomography and, magnetic resonance imaging amongst others. When all considered images in the registration problem are of the same modality, the registration is referred to as a monomodal registration. If at least two images of different modalities are to be registered, it is a multimodal registration. When performing intensity-based multimodal image registration, care should be taken to select suitable (dis)similarity measures (defined in the cost function, $\mathcal{C}$, in Eq 2.4) which can handle differences in the intensity distributions[1]. The most common examples of such measures are the (normalized) mutual information or the correlation ratio (Maes et al., 1997; Wells et al., 1996; Studholme et al., 1999; Roche et al., 1998a). Furthermore, multimodal imaging not only impacts the intensity distribution between two images of the same anatomical structure, it could also highlight physiological differences (such as cancerous tissue in PET imaging), significantly increasing the registration complexity.

In this work, we deal with a three image modalities: computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET).

[1]Note that this might even be necessary for monomodal registration. For example in the case of magnetic resonance images, where image intensities are typically not calibrated.

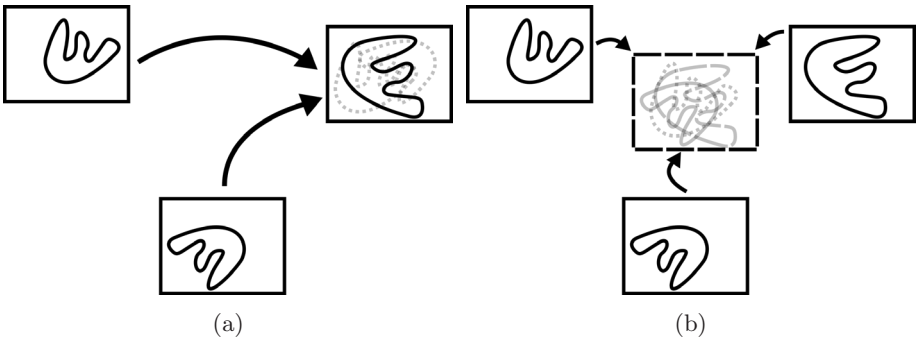


Figure 2.2: Registration of three images, employing two different transformation strategies. (a) All images are registered to the native space of one of the images in the set. (b) The images are registered to some artificial midspace.

2.2.5 Symmetric and asymmetric transformation strategies

Inherently, image registration as defined in Eq. 2.4 is an underdetermined problem. An infinite number of solutions for the set of transformations exist that align a set of images. For example, take one such solution and add a constant translation to every transformation in the set, resulting in a new set of transformations. This new set is equivalent to the initial set, as the same alignment is achieved. In practice, this degeneracy is broken either by setting one of the transformations in the set to zero (and registering all images to the native space of this image, referred to as an asymmetric strategy) or by forcing the transformations to some artificial or synthetic midspace (see Fig. 2.2). Several approaches can be followed to define this midspace (Balci et al., 2007; Metz et al., 2011; Joshi et al., 2004; Aganj et al., 2017) and are typically referred to as being symmetric, since the bias of choosing a 'reference' image to which all transformations are defined is removed. Part of this bias can be quantified in the transitivity of the transformations and will be discussed in later chapters. Additionally, symmetric registrations have been shown to be more accurate (Aganj et al., 2017).

Note that the choice of transformation strategy is independent of the number of images in the registration.

2.3 Pairwise and groupwise registration

The difference between pairwise and groupwise registration can be defined in two ways. Typically, it is defined based on the cardinality of the set of images for which alignment is needed. When the cardinality of that set of images is equal to two, the approach is typically referred to as a pairwise registration:

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{C}\left(I_1, I_2; \mathcal{T}_{\boldsymbol{\mu}_1}, \mathcal{T}_{\boldsymbol{\mu}_2}\right) \quad . \tag{2.5}$$

When three of more images are considered simultaneously, the approach is referred to as a groupwise registration. Note that in this work we only recognize approaches where the group of images in the registration is considered simultaneously in a global optimization as groupwise approaches. Approaches where three or more images are registered among each other in a pairwise manner are considered separately as *repeated* pairwise approaches.

Alternatively, it is possible to view groupwise registration as approaches for which the cardinality of the set of images does not need to be predefined and would theoretically work for any number of images. As such, pairwise registration can be studied as a special case of groupwise registration for which the cardinality of the set is predefined to be equal to two.

In recent years the increasing availability of imaging for an individual patient, for distinct pathologies and even across an entire population of patients has driven the interest in groupwise registration. In applications such as atlas construction (Joshi et al., 2004; Fletcher et al., 2009; Wu et al., 2011; Serag et al., 2012; Wu et al., 2016), multi-atlas based segmentation (Bhatia et al., 2007; Makropoulos et al., 2017), longitudinal population analysis (Davis et al., 2010; Huizinga et al., 2018) and spatiotemporal motion estimation (Metz et al., 2011; Yigitsoy et al., 2011; Royuela-del Val et al., 2016) groupwise registration is at the core of the methodological contribution. The unbiased estimation it can offer is most critical for population analyses, where an unbiased estimation of the population is essential, or in spatiotemporal motion estimation, where the simultaneous knowledge of multiple transformations restricts the registration to physically plausible solutions only.

2.4 Registration strategies for three or more images

We will consider three strategies to register three or more images. First, the simplest approach would be to select one image in the set as a privileged image and perform pairwise registrations to this privileged image with every other image in the set. The transformation between two arbitrary images can be found by inverting and composing the transformations obtained in the pairwise registrations. Note that such a strategy introduces a bias in the registration result, as the choice of the privileged image significantly impacts its behavior. This privileged image need not necessarily be the same as the reference image in the asymmetric transformation strategy defined in 2.2.5, but often is.

Secondly, when an implicit order of the images is available in the set (as is the case for dynamic acquisitions of the heart or lungs) it might be beneficial to register the first image to the second image, the second image to the third image and so on (in a pairwise manner). The transformation between two arbitrary images in the set is achieved by sequentially composing all intermediate transformations, potentially leading to compounding errors.

These first two methods can be considered as *repeated* pairwise approaches, as all transformations are optimized separately in a different pairwise registration. Herein, each pairwise registration only employs part of all image information in the group. Take as an example the registration of three hypothetical image modalities in Fig. 2.3. Registering image Fig. 2.3(a) and Fig. 2.3(b) would be impossible since they do not share any information that can be employed to register the images. However, including all images in a groupwise registration could provide enough common salient features such that the optimal alignment of Fig. 2.3(a) and (b) is achieved through aligning both of them with (c).

Where *repeated* pairwise approaches perform a set of registrations using two images in every registration, a groupwise approach only performs one registration using all images simultaneously. This offers two main advantages. First, bias is reduced since all transformations and images carry equal weight in the registration. Secondly, all images are employed simultaneously which could potentially aid the registration in terms of robustness and accuracy. Note that groupwise approaches are typically combined with symmetric transformation models discussed in section 2.2.5 to further reduce the bias in the registration algorithm.

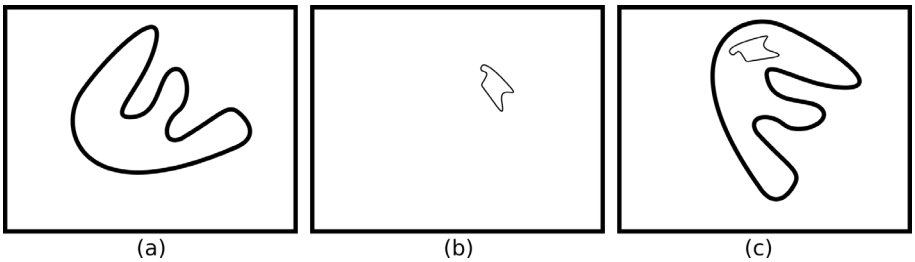


Figure 2.3: (a)-(c): Images of different modalities. Pairwise registration between (a) and (b) would be impossible. However, (c) might provide the additional information to properly align images (a) and (b) in a groupwise registration.

2.5 Situating this thesis

The work that was performed in the context of this thesis deals with advances in groupwise image registration. Groupwise registration has shown promise in a number of applications to increase registration accuracy, reliability or robustness, increased smoothness, reduced transitivity error and enhanced downstream analysis in the image processing pipeline. We compare groupwise registrations to repeated pairwise registration in several experiments and evaluate their differences in terms of registration accuracy and transitivity. Secondly, we fill a gap in current literature on efficient measures for multimodal groupwise registration.

These contributions are subdivided in four chapters. Thereafter we conclude with a general discussion and future perspectives.

In **Chapter 2** we investigate the added value of a symmetric (compared to an asymmetric) transformation strategy on the registration accuracy when registering two images. These experiments were performed specifically to clearly distinguish between both registration approaches controlling for all other possible variables such as the (dis)similarity measure and optimization strategies. A multimodal registration is performed in patients with oral squamous cell carcinoma where imaging, one computed tomography and one magnetic resonance image, was acquired of the head and neck region for surgical planning purposes.

For groupwise registration, an interesting class of (dis)similarity measures exists where the measure is expressed as the average of all pairwise comparisions between every image in the group and a template image. This allows the computational complexity of the algorithm to scale linearly with the number of images in the group. In **Chapter 3**, we investigate the influence of different

template images in groupwise similarity measures based on mutual information for monomodal data of the lungs (CT) and multimodal data of the brain (CT-MRI-PET). Additionally, an initial comparison is performed between repeated pairwise and groupwise registration based on the registration accuracy.

Building on the collected insights, we propose a novel groupwise similarity measure, the conditional template entropy, in **Chapter 4**. Herein, the similarity is measured as the average of the conditional entropy between every image in the group and a template image based on principal component analysis. Registrations were performed on monomodal data of the lungs (CT) and multimodal data of the brain (CT-MR-PET) and head and neck (MR, MR-CT). A comparison is carried out between repeated pairwise and groupwise registration approaches, with the registration accuracy and transitivity as the validation measures.

For huge groups of images, even a linearly scaling algorithm might be too computationally expensive, taking multiple days to perform a single registration. In **Chapter 5**, we focus on a scalable alternative and propose a novel dissimilarity measure based on Laplacian Eigenmaps, a non-linear dimensionality reduction technique. The proposed measure is compared against other state-of-the-art groupwise measures in terms of the registration accuracy and computation time in registrations performed on clinical multimodal data of the knee (MR) and brain (MR). Additionally, we perform experiments to investigate the effect the number of images has on the registration accuracy, transformation smoothness and computational time.

Finally, in **Chapter 6**, we discuss future research directions and summarize the findings in this thesis.

CHAPTER 3

Registration of magnetic resonance and computed tomography images in patients with oral squamous cell carcinoma for three-dimensional virtual planning of mandibular resection and reconstruction

Abstract

The aim of this study is to present and evaluate an automated method for registration of magnetic resonance imaging (MRI) and computed tomography (CT) or cone beam CT (CBCT) images in the mandibular region for patients with oral squamous cell carcinoma (OSCC). Registered MRI and (CB)CT could facilitate the 3D virtual planning of surgical guides employed for the resection and reconstruction of OSCC with mandibular invasion. MRI and (CB)CT images were collected retrospectively from 19 patients. MRI images were sequentially aligned with (CB)CT images employing a rigid registration approach (stage 1), a rigid registration approach using a mandibular mask (stage 2), and two non-rigid registration approaches (stage 3). Registration accuracy was quantified by the mean target registration error (mTRE), calculated over a set of landmarks annotated by two observers. Stage 2 achieved the best registration result with a mTRE of 2.5 ± 0.7mm, which was comparable to the inter- and intra-observer variabilities of landmark placement in MRI. Stage 2 was significantly better aligned compared to all approaches in stage 3. In conclusion, this study demonstrated that rigid registration with the use of a mask is an appropriate image registration method for aligning MRI and (CB)CT images of the mandibular region in patients with OSCC.

Based upon: Polfliet, M., Hendriks, M.S., Guyader, J.-M., ten Hove, I., Mast, H., Vandemeulebroucke, J., van der Lugt, A., Wolvius, E.B., & Klein, S. (2021). Registration of magnetic resonance and computed tomography images in patients with oral squamous cell carcinoma for three-dimensional virtual planning of mandibular resection and reconstruction. *International Journal of Oral and Maxillofacial Surgery.*.

3.1 Introduction

Oral squamous cell carcinoma (OSCC) is the sixth most common cancer worldwide. The overall 5-year survival rate of patients with OSCC is less than 50% and has not shown significant improvement over the last decades, despite advances in treatment modalities (Binahmed et al., 2007; Carvalho et al., 2005; Nieberler et al., 2016; Shah and Gil, 2009; Safi et al., 2017).

Segmental mandibular resection and reconstruction with a free vascularized osseocutaneous flap is currently the recommended treatment for OSCC invading the mandible (Succo et al., 2015). The main goal of surgical treatment is to obtain tumor free resection margins, with acceptable remaining function (chewing, swallowing and speaking) and physical appearance. Achieving tumor free resection margins is challenging, but crucial for disease control and survival (Binahmed et al., 2007; Nieberler et al., 2016; Smits et al., 2018; Kademani et al., 2005; Kreppel et al., 2011; Barroso et al., 2018; Dillon et al., 2015; Varvares et al., 2015). Recently, inadequate resection margins were found in 20% of the bone resections which negatively impacted the 5-year survival rate of patients (Binahmed et al., 2007; Nieberler et al., 2016; Smits et al., 2018). The inability to intra-operatively distinguish tumor from healthy bone tissue during resection is the most common cause for such inadequate margins (Barroso et al., 2018). Additionally, a re-resection of positive margins in a second operation is not desirable, due to technical difficulties and has a negative effect on the survival of the transplant(Dillon et al., 2015). During resection, tumor free surgical margins are the only prognostic factor that the surgeon can control(Dillon et al., 2015).

The state-of-the-art mandibular reconstruction method is based on preoperative three-dimensional (3D) virtual surgical planning using 3D printed surgical guides (Varvares et al., 2015; Tarsitano et al., 2015; Cornelius et al., 2016). Herein, the patient undergoes the necessary imaging, after which the surgeon virtually defines the cutting planes and plans the resection and subsequent reconstruction. Thereafter, the surgical guides are printed and the virtual planning is translated to the surgical procedure. Mandibular and fibular cutting guides have shown to provide a better fit of the fibula parts resulting in reduction of surgical time. However, accurate 3D virtual planning of the surgical cutting guides remains essential in order to achieve complete resection.

Current 3D virtual planning is based on computed tomography (CT) or cone beam CT (CBCT) images, which offer detailed information on bone geometry and cortical bone destruction but do not provide accurate information on bone marrow involvement and perineural spread of the tumor. In recent years, mag-

netic resonance imaging (MRI) is increasingly used for diagnostic purposes due to its better visualization of tumor tissue, mandibular bone marrow involvement and perineural spread along the interior alveolar nerve (Nieberler et al., 2016; Li et al., 2014a,b; Blatt et al., 2016).

The uncertainty about the location of tumor boundaries in the 3D virtual planning based on (CB)CT acquisitions could be eliminated by including MRI aquisitions (Dong et al., 2011). Overlaying or fusing these MRI and (CB)CT images before the preoperative virtual planning could aid the performing surgeon in defining the surgical guides and subsequent reconstruction thanks to a more accurate determination of the osteotomy location and a better understanding of the surrounding structures. The integration of fused CT and MRI imaging in clinical practice of mandibular resection planning has been shown to be a safe and accurate alternative (Kraeima et al., 2018). However, due to a different orientation and position of the mandible in the MRI and (CB)CT acquisitions, an image registration method is required to establish the spatial correspondences between the different images.

The aim of this work is to present and evaluate an automated method to perform image registration of MRI and (CB)CT in the mandibular region in patients with oral squamous cell carcinoma, which could subsequently be integrated in a pipeline for virtually planning of mandibular resections and reconstructions.

3.2 Materials and methods

3.2.1 Dataset

The study was reviewed and approved by the local medical ethics review committee (MEC-2016-143), and performed in accordance with national and international legislation. The need for informed consent was waived owing to the retrospective and anonymized nature of the study. Preoperative 3D MRI and (CB)CT scans of the head-and-neck region were collected retrospectively from 19 patients diagnosed between 2014 and 2016 with untreated primary OSCC with invasion of the mandible. The images were anonymized prior to processing. The MRI scans were acquired with a Spin Echo T1-weighted sequence. The in-plane voxel size of MRI was between 0.4 x 0.4 and 0.5 x 0.5 mm^2 and slice thickness was between 3 and 4 mm. Echo time (TE) ranged from 10.8 to 13.6 ms, the repetition time (TR) from 416 to 689 ms, and the flip angle (FA) was 90, 111 or 160 degrees. CT imaging in-plane voxel size ranged from 0.3

x 0.3 mm^2 to 0.5 x 0.5 mm^2 and slice thickness from 0.3 to 0.6 mm. CBCT imaging in-plane voxel size was 0.3 x 0.3 mm^2 and slice thickness was 1 mm. The mean time between the MRI and (CB)CT scans was 9 days (2-23 days, SD: 5.3). No pre- or post-processing was applied to the images and they remained unmodified for the registration.

3.2.2 Registration

We investigated an automated image registration method to align the MRI with the (CB)CT images; herein the alignment is achieved in three stages. In the first stage, an initial rigid alignment was estimated. Thereafter a more refined rigid alignment was estimated focused around the mandible. In the third and final stage, a deformable alignment was performed, for which different approaches were compared. In all stages we used an automated intensity-based 3D registration framework (elastix) (Klein et al., 2010) based on the maximization of mutual information (Thévenaz and Unser, 2000) using a stochastic gradient descent optimization method (Klein et al., 2009b).

In the first stage, two consecutive registrations were performed to achieve an initial rigid alignment. First, a global translation was estimated, since the MRI images needed to be roughly aligned to the (CB)CT image domain. Subsequently, a rigid registration was carried out, estimating both translations and rotations (parameterized by Euler angles).

In the second stage, the initial rigid alignment was fine-tuned by restricting the focus of the algorithm on a 3D region of interest encompassing the mandible, manually drawn in the (CB)CT image. As such, all image information outside of the region of interest is ignored and potential registration difficulties due to pose or appearance changes could be alleviated. The mandibular mask was drawn slice by slice, using open-source ITK-SNAP software (Yushkevich et al., 2006).

In the third stage, we evaluated whether the alignment could be refined further using a non-rigid (or deformable) registration to compensate for any geometric distortions in the MRI images (Chang and Fitzpatrick, 1992). A parametric B-spline free-form deformation model was employed, and (isotropic) control point spacings of 64, 32, and 16 mm were evaluated (Rueckert et al., 1999). Furthermore, two different approaches for the deformable registration were compared: a) an asymmetric approach with the (CB)CT image as the fixed (or target, reference) image and the MRI as the moving (or template, source) image; and b) a symmetric approach were both images were registered to a common mid-space (Bhatia et al., 2004). Results from literature suggest

that symmetric registration techniques can lead to improved registration accuracy and inverse-consistency (Aganj et al., 2017; Lorenzi et al., 2013), which are especially critical for treatment planning (Rivest-Hénault et al., 2015). All stages are illustrated in Fig. 3.1. The three registration stages are executed in a consecutive manner. The registration result after stage 1 was employed to initialize the registration in stage 2. The registration result after stage 2 was employed to initialize the registration in stage 3, where either approach 3a or approach 3b was taken.

3.2.3 Evaluation

Registration accuracy in the mandibular region was evaluated in terms of mean target registration error (mTRE), by computing the Euclidean distance between corresponding landmarks in MRI and (CB)CT and then averaging it over all landmarks (Fitzpatrick and West, 2001). An extensive landmark set was designed with 39 anatomical reference points in order to evaluate the registration error for the entire mandible specifically. The set consisted of 22 landmarks placed at the roots of each lower tooth (for the molars, both roots were considered as landmarks) and 17 anatomical reference points on the mandible. The set of landmarks is described in detail in Table 3.1 and Fig. 3.2. Due to tumor invasion in the bone or removed tooth elements, not all landmarks from the dataset could be annotated for all images.

The landmarks were annotated manually in each MRI and (CB)CT acquisition by two researchers (M.S.H., J.-M.G.) who were trained in advance. The reliability of each landmark was investigated by calculating the interobserver variability for both MRI and (CB)CT images separately. In addition, one observer (M.S.H.) repeated the annotation of all landmarks, to enable the assessment of the intraobserver variability. Intra- and interobserver variability were quantified by calculating the Euclidean distance (similar to the mTRE) between corresponding landmarks of the same and different observer, respectively.

If the interobserver variability for a landmark was greater than 5 mm, the landmark was excluded (when applicable both for left and right sides). After exclusion of the unreliable landmarks, the mTRE was calculated for each registration stage to assess the registration accuracy.

3.2.4 Statistical analysis

Two statistical analyses were performed. First, a comparison was performed between the stage that performed best (with the lowest mTRE) and all other

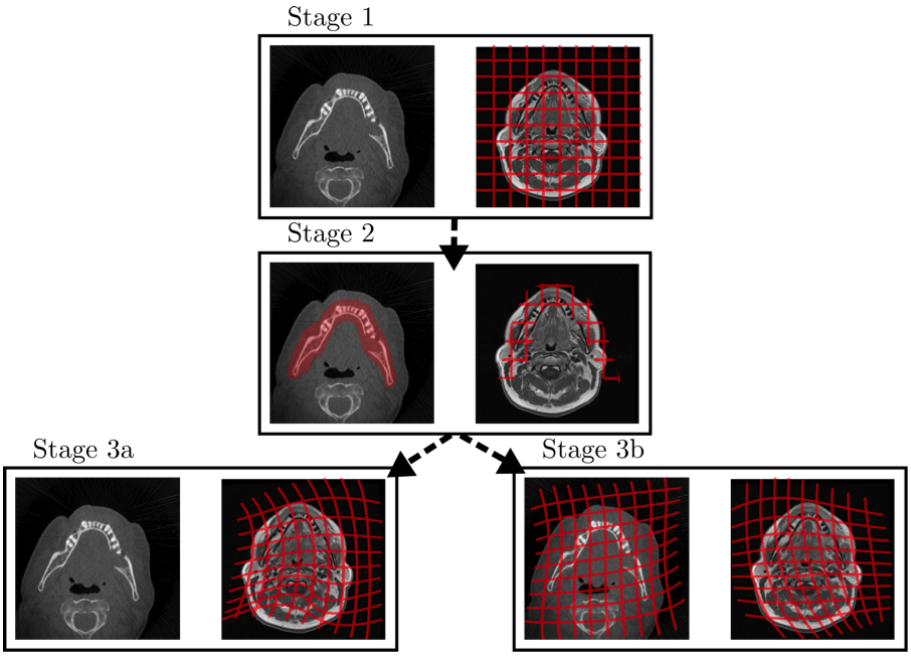


Figure 3.1: The proposed three-stage registration method to align MRI with (CB)CT images in the mandibular region. In the first stage, an initial rigid alignment was estimated. Thereafter, in the second stage, a more refined rigid alignment focused around the mandible was estimated. In the third (and final) stage two approaches for deformable alignment were investigated side-by-side, an asymmetric (3a) and a symmetric approach (3b). Note that, unlike the schematic 2D illustrations in this figure, all registrations were performed completely in the 3D space.

Table 3.1: Description of the anatomical landmarks on the teeth and the mandible.

Abbreviation	Landmark	Description
t31, t32, t33, t34, t35, t36, t37, t38, t41, t42, t43, t44, t45, t46, t47, t48	Teeth of the third and fourth quadrant.	22 landmarks of the teeth; the molars (36, 37, 38, 46, 47, 48) have two roots.
A	Menthon	The most inferior point of the mandibular symphysis.
B	Mental Foramen	Foramen located on the anterior side of the mandible.
C	Gonion	A point defined as the mandibular angle, representing the intersection of the lines of the posterior ramus and the inferior border of the mandible.
D	Mandibular foramen	Foramen located on the internal surface on the ramus.
E	Coronoid process	The tip of the coronoid process
F	Left Condylion	Leftmost aspect of the condylar head.
G	Right Condylion	Rightmost aspect of the condylar head.
H	Top Condylion	Top of the condylar head.
I	Mandibular notch	Notch located at the most superior point of the ramus, which separates the coronoid process anteriorly and the condyloid process posteriorly.

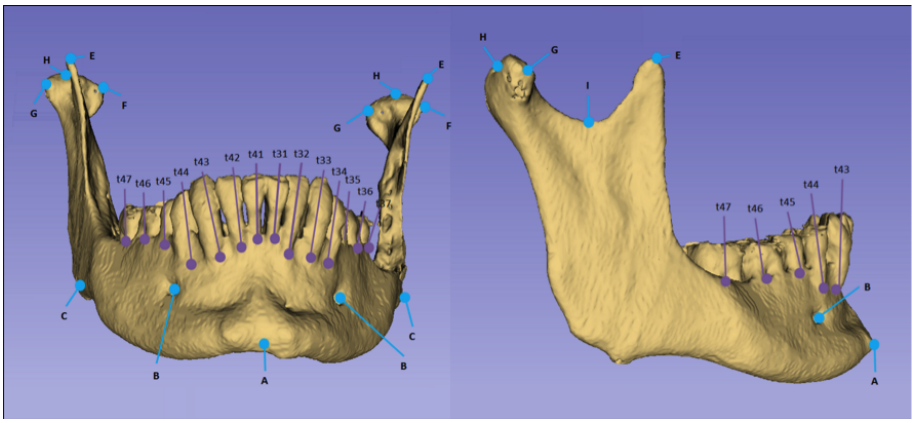


Figure 3.2: Tooth and mandibular evaluation landmarks. The purple and the blue landmarks indicate the apex of the teeth and the anatomical positions of the mandible, respectively. These 39 landmarks correspond with the landmarks described in Table 3.1. Landmark D is not illustrated in this figure, because this landmark is behind the field-of-view. Note that the 3D model shown here was only generated for illustration purposes; during annotation of the landmarks, the original (CB)CT and MR acquisitions were used, inspecting image slices in three orthogonal planes (axial, sagittal, coronal).

registration stages. Secondly, in the third stage a comparison was performed between the asymmetric and symmetric approach for each spacing of the control points (64mm, 32mm and 16 mm). A two-sided Wilcoxon signed rank test at a significance level of 0.05 was used to evaluate the differences in the distribution of the mTRE.

3.3 Results

The interobserver variability for the landmarks Gonion (C) and Coronoideus (I) was greater than 5 mm (6.5 mm and 6.0 mm, respectively). These landmarks on both sides of the mandible were therefore excluded in the calculation of the mTRE and the observer variabilities. On average, 20 landmarks per patient remained to calculate the mTRE.

The inter- and intraobserver variability (mean±SD over all subjects) for landmark placement were found to be 2.4 ± 0.7 mm and 2.0 ± 0.5 mm, respec-

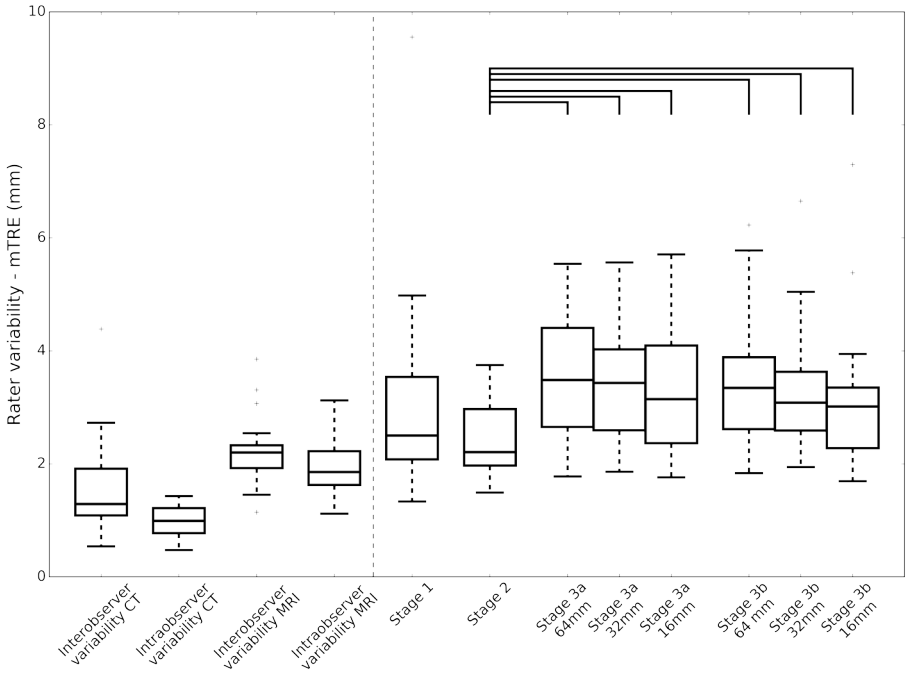


Figure 3.3: Boxplots representing the distributions of the inter- and intraobserver variabilities for CT and MRI and of the mTRE values for each registration stage. The square brackets between two connected boxplots indicate statistical significance at a level of 0.05.

tively, for MR images and 1.5 ± 0.8 mm and 1.0 ± 0.3 mm for (CB)CT images.

After rigid registration in stage 1, the mTRE (mean±SD over all subjects) was 3.1 ± 1.8 mm. After rigid registration with the use of a mask around the mandible (stage 2), the mTRE was 2.5 ± 0.7 mm. After asymmetric non-rigid registration (stage 3a) with B-spline control point spacings of 64, 32, and 16 mm, the mTREs were 3.6 ± 1.1 mm, 3.4 ± 1.1 mm, and 3.3 ± 1.2 mm, respectively. In the symmetric non-rigid registration (stage 3b) the mTRE values were found to be 3.5 ± 1.2 mm, 3.3 ± 1.2 mm, and 3.1 ± 1.3 mm. Fig. 3.3 shows the distributions of mTRE over all subjects for each registration stage, as well as the inter- and intraobserver variability.

Compared to stage 2, which yielded the lowest average mTRE, stage 3

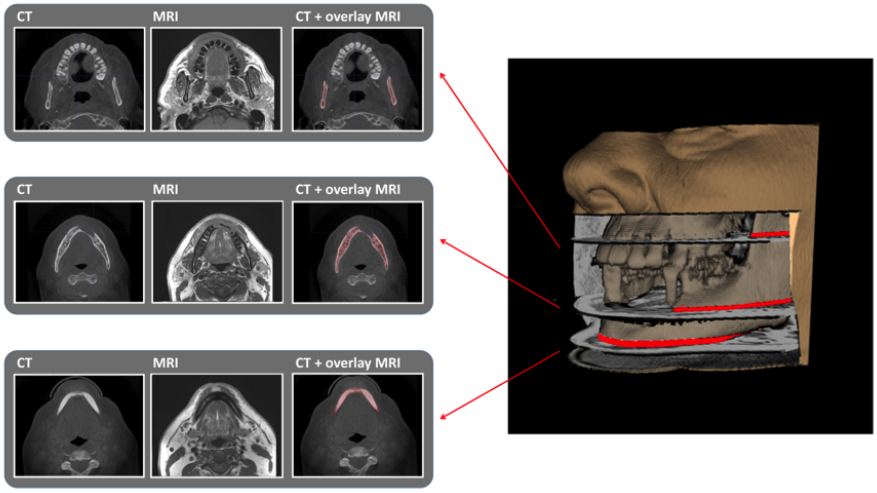


Figure 3.4: This figure illustrates the result of the registration of MRI and CT images after stage 2 of a representative case (subject 8). In three randomly selected slices, the mandible was segmented in the registered MR images and overlaid in red with the corresponding CT images. These results were combined in a 3D rendering with the segmented skull and mandible from the CT images. The gap in the front represents tumor tissue which is invading the mandible in and around the anterior part of the mandible.

produced significantly different mTRE values (3a, 64mm: W=5, $p< 0.001$; 3a, 32mm: W=2, $p< 0.001$; 3a, 16mm: W=6, $p< 0.001$; 3b, 64mm: W=3, $p< 0.001$; 3b, 32mm: W=15, $p= 0.001$; 3b, 16mm: W=26, $p= 0.005$). No significant differences were found between stage 1 and stage 2 (W=47, $p= 0.054$) or between stage 3a and stage 3b (64mm: W=59, $p= 0.147$; 32mm: W=69, $p= 0.294$; 16mm: W=48, $p= 0.059$). A representative case illustrates the result of the registration of MRI and CT images after stage 2 in Fig. 3.4.

Regarding computation time, stage 1 required 266 ± 65 s to complete, and stage 2 required an additional 131 ± 37 s. Stage 3a required an additional 285 ± 48 s, 288 ± 34 s and 294 ± 51 s, for 64mm, 32mm and 16mm, respectively, whereas in stage 3b those computations required 1138 ± 140 s, 1234 ± 139 s

and 1356 ± 151 s. All experiments were performed single-threaded on the local university CPU cluster.

3.4 Discussion

Our study shows that the rigid registration with a mask (stage 2) is the recommended method for registering MRI and (CB)CT images in the mandibular region. Stage 2 achieved a lower mTRE compared to stage 1 although the difference was not significant (p=0.054). However, the localized focus in stage 2 should generalize better to other patients and be more robust to outliers. We have provided a protocol for applying the recommended method in the appendix.

In this work two approaches for deformable registration were applied. A conventional asymmetric approach where the (CB)CT image was employed as the fixed image and the MRI image as the moving image (stage 3a) and a symmetric approach where the images were registered to a common reference system (stage 3b). Although the differences were not statistically significant, stage 3b achieved a marginal improvement compared to stage 3a for all control point spacings of 64, 32 and 16 mm. Furthermore, the registration error of all approaches in stage 3 was significantly higher than the error in stage 2. As such, our results indicate that non-rigid registration (stage 3) has no added value in this application. Thanks to the large number of landmarks (35 after exclusion), the mTRE could be reliably estimated. This was even the case for patients for which not all landmarks could be annotated, e.g. cases where tooth elements were extracted, cases with bone invasion by tumor, and cases with a landmark outside the field-of-view. The inter- and intraobserver variability of landmark annotations in (CB)CT images were consistent with those reported in literature (Ludlow et al., 2009; Lagravère et al., 2010). Furthermore, the mTRE achieved by the best registration method (stage 2: 2.5 mm) was similar to the inter- and intraobserver variabilities for MRI images (2.4 mm and 2.0 mm, respectively). As such, lower mTRE values based on the landmarks employed in this study can hardly be expected. Note that a recent landmark accuracy study in MR images found a lower inter- and intraobserver variability (Juerchott et al., 2020). However, this difference can be explained by the considerably lower slice thickness (0.53 mm vs. at least 3mm in our study) employed therein. Based on the inter- and intraobserver variabilities, the Gonion and Coronoideus landmarks were excluded; their annotations most likely were hindered by the poor delineation and fuzzy boundaries of the landmark location (Williams and Richtsmeier, 2003). Note that several other evaluation

methodologies for registration accuracy exist, such as the Euclidean distance between centroids, overlap measures and surface distances of manually segmented anatomical structures (Murphy et al., 2011a; Rohlfing, 2011). Here, we opted for manual landmark annotations, since a relatively large number of well-defined landmarks could be identified, allowing a reliable estimation of mTRE, while manual segmentation would have been much more time-consuming.

To our knowledge, data regarding registration of MRI and (CB)CT images of the head-and-neck region are scarce. Previous studies (Kraeima et al., 2015; Fortunati et al., 2014; Leibfarth et al., 2013; du Bois dAische et al., 2007; Webster et al., 2009) found registration errors in the range of 1.7-3.3 mm in datasets of 4-16 patients. None of these studies have focused specifically on the mandible, hindering a thorough comparison with our study. However, the results of these studies suggest that the achieved mTRE of 2.5 mm (for stage 2) indicates a state-of-the-art error level.

Fortunati et al. (2015) suggested that patient immobilization during imaging leads to better registration of the MRI and (CB)CT images of the head-and-neck region. Their study found a registration error of 7.0 mm without immobilization and a registration error of 1.9 mm with immobilization. Implementation of immobilization equipment in our specific application might not add value in clinical practice, since a competitive mTRE of 2.5 mm was already achieved without immobilization. Moreover, rigorous immobilization of the mandible would be challenging, and likely not comfortable for the patient. The mandibular mask used in stage 2 was drawn manually around the mandible in each slice of the (CB)CT images. Although the mask does not have to be delineated very precisely (it just serves to indicate an approximate region of interest), this manual interaction step may not be desirable in clinical practice. Development of a robust semi-automated or even fully automated segmentation (Egger et al., 2018) is therefore recommended to accelerate this step. We refer the reader to a recent review of such methodologies for a full overview (Wallner et al., 2019).

Although we used the open-source Elastix software to implement the image registrations in this study, other (open-source or commercial) softwares that implement similar registration algorithms based on maximization of mutual information could have been used as well. Some well-known open-source examples include NiftyReg (Modat et al., 2010) and ITKv4/ANTS (Avants et al., 2014). After proper configuration, these tools are expected to achieve similar registration accuracy.

Correct determination of the osteotomy location depends on several factors (e.g. the waiting time between imaging and surgery, the process of the translation from 3D virtual planning to the patient, the accurate placement of the

cutting guide during surgery, the length of the fibula reconstruction, relation of the the tumor to the mental nerve and the remaining teeth), which could independently contribute to positive resection margins. For example, several studies have shown that CBCT results in less accurate 3D planning models than CT (van Baar et al., 2018). When the proposed registration method would be translated into clinical practice, image registration errors have to be considered in conjunction with all other sources of errors. In this work no clinical outcome criteria such as the resection margin, the frequency of revising the surgical guide during planning or frequency of local tumor progression were employed as it entails a retrospective study. A randomized clinical trial is needed to reveal the added value of registered MRI and CB(CT) in virtual planning of mandibular resection and reconstruction.

3.5 Conclusion

This study presented an image registration method for aligning MRI and (CB)CT images of the mandibular region in patients with OSCC. We showed that rigid registration within a region of interest drawn around the mandible is the recommended registration method for the alignment of MRI and (CB)CT images in the mandibular region.

Acknowledgement

This study was partially funded by Fonds NutsOhra (nr. 1404-0480). The authors would also like to thank the anonymous reviewers for their feedback and remarks on this work, which substantially improved its quality.

CHAPTER 4

The Pythagorean averages as template images in efficient groupwise registration

Abstract

Many applications in medical image processing can benefit from robust and unbiased groupwise registration. However, no obvious solution is available for multimodal registration problems involving a large number of images. A technique that is frequently applied calculates the sum of the pairwise similarities between a template image and all the images in the group. This allows the algorithm to scale linearly with respect to the number of images involved. Typically the arithmetic average is used as the template image, which has been shown to be a poor choice. We present geometric and harmonic averaging as an alternative and validate their performance in experiments on intrasubject dynamic acquistions of the lung (CT) as a monomodal use-case and intrasubject acquisitions of the brain CT-MR-PET) as a multimodal use-case. These experiments show an increased robustness and accuracy compared to the arithmetic average.

Based upon: Polfliet, M., Klein, S., Huizinga, W., De Mey, J., & Vandemeulebroucke, J. (2016). The Pythagorean averages as group images in efficient groupwise registration. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)* (pp. 1261-1264). IEEE.

4.1 Introduction

Many applications in medical image processing can benefit from robust and unbiased groupwise registration. Population analyses, motion estimation across temporal sequences or perfusion analysis in dynamic contrast-enhanced sequences are some examples. With the possibility of a large number of images present in these sequences, scalability of the approach is an important property. For multimodal images, efficient and scalable (dis)similarity measures are lacking.

Applying entropy-based measures, such as mutual information, directly to groupwise registration leads to a joint probability space whose size increases exponentially with respect to the number of images involved. Its application is limited by the sparsity herein, effectively suffering from the *curse of dimensionality*. Several authors have tried solving this by approximating the probability density functions (PDFs), but eventually falling to sparsity as well when more images were added.

An interesting class of (dis)similarity measures iteratively calculates a group or template image and determines the global (dis)similarity as the sum of the pairwise (dis)similarities between every image in the group and this template image (Bhatia et al., 2007). This leads to an approach with a linearly increasing computational complexity with respect to the number of images in the group, thereby making its application to problems with a large number of images feasible. The template image is typically chosen to be the arithmetic average image.

However, arithmetic averaging in the context of multimodal data is not optimal, given the potential scale and range differences in intensities. Furthermore, it has been shown that the sharpness of the template image can have an impact on the robustness and accuracy of the registration (Wu et al., 2011).

In this work we investigate two novel similarity measures for multimodal groupwise registration in which the geometric and harmonic averages are calculated as the template images. The proposed measures evaluated in intrasubject dynamic acquistions of the lung (CT) as a monomodal use-case and intrasubject acquisitions of the brain CT-MR-PET) as a multimodal use-case and compared to other state-of-the-art measures.

4.2 Materials and methods

4.2.1 Groupwise registration

Groupwise registration can be defined as an optimization problem of n transformations which align n images in a group such that

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{C}\left(I_1, \ldots, I_n; \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, \mathcal{T}_{\boldsymbol{\mu}_n}\right) \quad . \tag{4.1}$$

Herein is $\mathcal{T}_{\boldsymbol{\mu}_i} : \Omega \subset \mathbb{R}^d \to \Omega_i \subset \mathbb{R}^d$ the transformation that maps the coordinates from the common reference domain, Ω, to the domain of the ith image, Ω_i, and $\boldsymbol{\mu}_i$ are the parameters that define $\mathcal{T}_{\boldsymbol{\mu}_i}$. For the optimization, $\boldsymbol{\mu}$ is defined as a vector formed by the concatenation of all separate transformation parameters, $\boldsymbol{\mu}_i$. $I_i : \Omega_i \subset \mathbb{R}^d \to \mathbb{R}$ is the continuous intensity function associated to the ith image for which we assumed an interpolation scheme and $\mathcal{C}$ the cost function or objective value of the registration problem.

The cost function $\mathcal{C}$ is commonly given by a weighted sum of a (dis)similarity measure, $\mathcal{S}$ or $\mathcal{D}$, and a regularizer $\mathcal{R}$

$$\mathcal{C} = \mathcal{D} + \lambda \mathcal{R} \quad , \tag{4.2}$$

where λ is the weight associated to the regularizer.

4.2.2 Registration measures

Similar to Bhatia et al. (2007) we extended pairwise mutual information, $\mathcal{S}_{MI}$, to the arithmetic average mutual information

$$\mathcal{S}_{AAMI}\left(I_1, \ldots, I_n; \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, \mathcal{T}_{\boldsymbol{\mu}_n}\right) = \frac{1}{n} \sum_{i=1}^{n} \mathcal{S}_{MI}\left(I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}, \overline{I}^A_{\boldsymbol{\mu}}\right) \quad , \tag{4.3}$$

with $\overline{I}^A_{\boldsymbol{\mu}}$ the voxel-wise arithmetic average intensity image as the template image defined as

$$\overline{I}^A_{\boldsymbol{\mu}}\left(\boldsymbol{x}\right) = \frac{1}{n} \sum_{i=1}^{n} I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}\left(\boldsymbol{x}\right) \tag{4.4}$$

and $\mathcal{S}_{MI}$ the pairwise mutual information and $\circ$ refers to the composition of two functions.

Note that an analogous measure for mono-modal registration was previously proposed, the sample variance (SV)(Bhatia et al., 2007), and is in essence the groupwise extension of the pairwise mean squared differences (MSD) measure

$$\mathcal{D}_{SV}\left(I_1,\ldots,I_n;\mathcal{T}_{\boldsymbol{\mu}_1},\ldots,\mathcal{T}_{\boldsymbol{\mu}_n}\right)=\frac{1}{n}\sum_{i=1}^{n}\mathcal{D}_{MSD}\left(I_i\circ\mathcal{T}_{\boldsymbol{\mu}_i},\overline{I}^{A}_{\boldsymbol{\mu}}\right)\quad, \tag{4.5}$$

Wu et al. (2011) demonstrated that sharpness of the template image is of great importance for $\mathcal{D}_{SV}$. The influence of the sharpness of the image has not been investigated for multimodal groupwise registration derived from mutual information. Consider the pairwise mutual information $\mathcal{S}_{MI}$ in Eq. (4.3) defined as

$$\mathcal{S}_{MI}\left(I_i\circ\mathcal{T}_{\boldsymbol{\mu}_i},\overline{I}_{\boldsymbol{\mu}}\right)=H\left(I_i\circ\mathcal{T}_{\boldsymbol{\mu}_i}\right)+H\left(\overline{I}_{\boldsymbol{\mu}}\right)-H\left(I_i\circ\mathcal{T}_{\boldsymbol{\mu}_i},\overline{I}_{\boldsymbol{\mu}}\right)\quad, \tag{4.6}$$

where H is the (joint) entropy of the image(s). The influence of the sharpness of the template image can be intuitively understood considering the toy example in Fig. 4.1. Eleven black and white images (Fig. 4.1(a)) are progressively shifted along the horizontal axis. Contrary to pairwise registration where the entropy of the two images stays relatively constant, in groupwise registration the entropy of the template image can vary significantly. Closer to alignment, the sharpness of the template image is typically increased which leads to a decrease in the entropy of the template image as can be seen in Fig. 4.2(c). The behavior of this term is counter productive and can challenge the convergence of the optimization by effectively cancelling out the positive contribution of the joint entropy to the global mutual information measure.

Additionally, when dealing with multimodal registration problems it is common that images exhibit different intensity ranges and scales which arithmetic averaging handles very poorly. This effect may be reduced through preprocessing techniques such as normalizing the intensities between 0 and 1, using z-scores or histogram equalization.

In this work we will investigate the voxel-wise geometric and harmonic average as template images, instead of the arithmetic average, given as

$$\overline{I}^{G}_{\boldsymbol{\mu}}\left(\boldsymbol{x}\right)=\sqrt[n]{\prod_{i=1}^{n}I'_i\circ\mathcal{T}_{\boldsymbol{\mu}_i}\left(\boldsymbol{x}\right)} \tag{4.7}$$

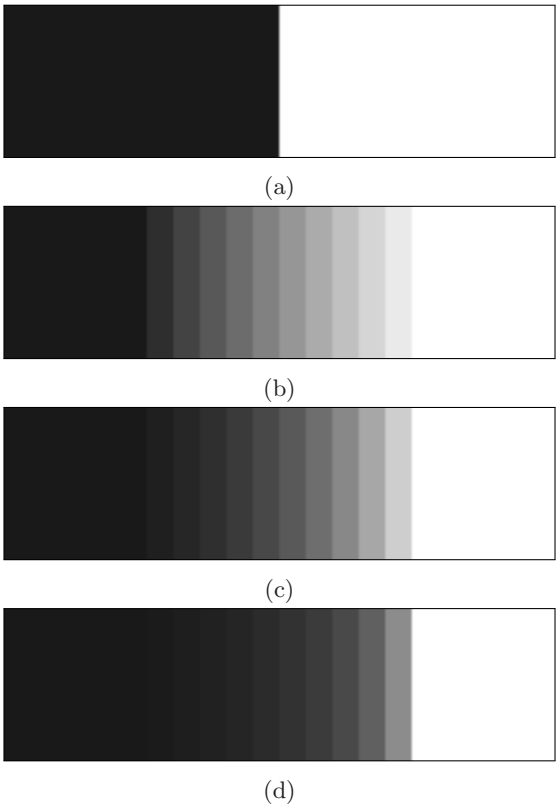


Figure 4.1: Toy example in which eleven black-and-white images are initially misaligned. (a) A single black-and-white image. (b) The arithmetic average image, (f) The geometric average image and (g) The harmonic average image at maximal misalignment.

$$\overline{I}^{H}_{\boldsymbol{\mu}}(\boldsymbol{x}) = \frac{n}{\sum\limits_{i=1}^{n} \frac{1}{I'_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}(\boldsymbol{x})}} = \frac{n \prod\limits_{i=1}^{n} I'_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}(\boldsymbol{x})}{\sum\limits_{i=1}^{n} \prod\limits_{j=1, j\neq i}^{n} I'_j \circ \mathcal{T}_{\boldsymbol{\mu}_j}(\boldsymbol{x})} \quad . \tag{4.8}$$

and their utility when used as the geometric and harmonic average mutual

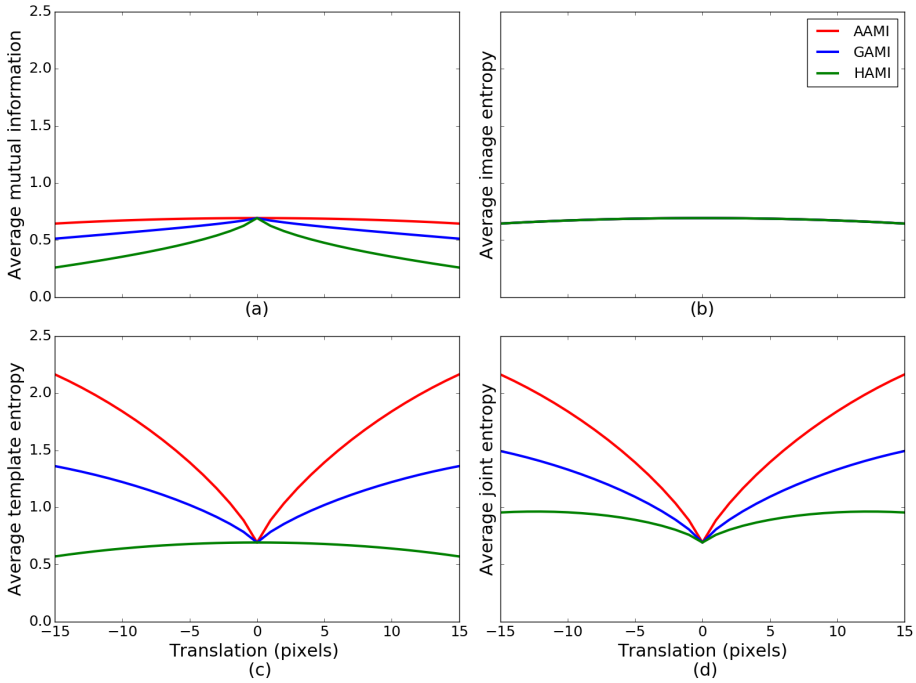


Figure 4.2: Toy example with eleven black-and-white images. (a) The value of the average mutual information, (b) the average entropy of all images in the group, (c) the average entropy of the template image and (d) the average joint entropy of the images in the group with the template image for different translations.

information (GAMI, HAMI)

$$\mathcal{S}_{GAMI}\left(I_1 \circ \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, I_n \circ \mathcal{T}_{\boldsymbol{\mu}_n}\right) = \sum_{i=1}^{n} \mathcal{S}_{MI}\left(I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}, \overline{I}_{\boldsymbol{\mu}}^{G}\right) \tag{4.9}$$

$$\mathcal{S}_{HAMI}\left(I_1 \circ \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, I_n \circ \mathcal{T}_{\boldsymbol{\mu}_n}\right) = \sum_{i=1}^{n} \mathcal{S}_{MI}\left(I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}, \overline{I}_{\boldsymbol{\mu}}^{H}\right) \tag{4.10}$$

The multiplication in Eq. (4.7) and (4.8) handles range and scale differences better than the summation in Eq. (4.4). I_i' was obtained using a linear intensity mapping to avoid negative values. Additionally, since the geometric and

harmonic averages accentuate low intensities they are more specific and only structures in overlap in all images will be represented in the template image. This specificity leads to a sharper template image. GAMI and HAMI show the desired optimization behavior in Fig. 4.2.

The proposed measures were implemented in the registration packages `elastix` and did not introduce any computational overhead compared to arithmetic averaging. The registrations were performed with an adaptive stochastic gradient descent and multiresolution scheme (Klein et al., 2009b).

4.2.3 Thoracic 4D CT

As a monomodal test-case we applied the proposed similarity measures to 4D-CT data which were taken from the POPI and DIR-LAB datasets (Vandemeulebroucke et al., 2011; Castillo et al., 2009)(Fig. 4.3). They contained respectively 6 and 10 patients with ground-truth annotations of anatomical landmarks. The quality of the registration was expressed as the groupwise target registration error (gTRE) with respect to the first image in the sequence

$$\mathrm{gTRE}\left(\boldsymbol{\mu}\right) = \frac{1}{n}\sum_{i\neq r}^{n}\frac{1}{|P_i|}\sum_{j}^{|P_i|}||\mathcal{T}_{i,r}\left(\boldsymbol{p}_{i,j}\right) - \boldsymbol{p}_{r,j}|| \tag{4.11}$$

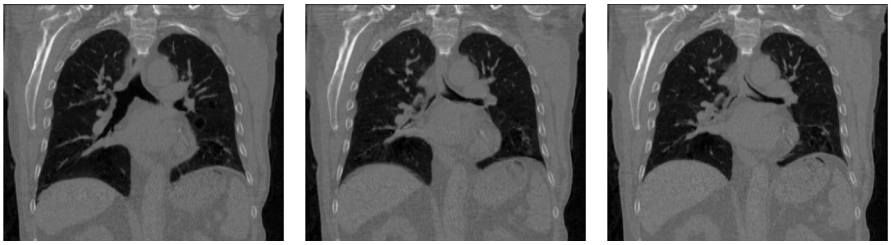
Herein is r the index of the reference image for the calculation of the gTRE, P_i the set of landmarks in the ith image and $|P_i|$ is the cardinality of that set. $\mathcal{T}_{i,r}$ is the transformation that maps the coordinates from the ith image to the reference image and $\boldsymbol{p}_{i,j}$ the jth landmark from the ith image. $||\cdot||$ refers to the length of the enclosed vector.

A deformable transformation model was used to align the images based on cubic B-splines with a final control point spacing of 12.0 mm. Lung masks were obtained as described by Vandemeulebroucke et al. (2011). We compared a total of two pairwise (dis)similarity measures ($\mathcal{D}_{MSD}$ and $\mathcal{S}_{MI}$) and four groupwise measures ($\mathcal{D}_{SV}$, $\mathcal{S}_{AAMI}$, $\mathcal{S}_{GAMI}$ and $\mathcal{S}_{HAMI}$). The pairwise registrations were performed with the first image in the group (the inspiration image) as a reference. As a regularizer we used the pairwise and groupwise bending energy (GBE)

$$\mathcal{R}_{GBE}\left(\mathcal{T}_{\boldsymbol{\mu}_1},\ldots,\mathcal{T}_{\boldsymbol{\mu}_n}\right) = \frac{1}{|S|}\sum_{\boldsymbol{x}\in S}\frac{1}{n}\sum_{i=1}^{n}\sum_{l,m=1}^{d}\left(\frac{\partial^2\mathcal{T}_{\boldsymbol{\mu}_i}(\boldsymbol{x})}{\partial x_l \partial x_m}\right)^2 . \tag{4.12}$$

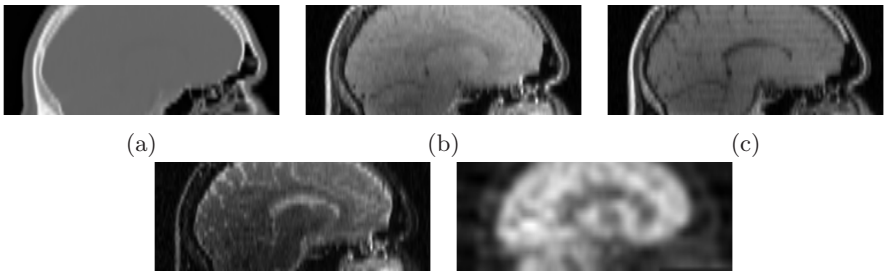
Herein is d the dimension of the images and $\boldsymbol{x}$ the coordinate samples take from the set, S, which was drawn from the image domain and $|S|$ is the cardinality

of that set. The weights for the regularizer were determined empirically to best fit each specific metric, noise level and transformation. Note that it is possible to start this empirical search at a best-guess value dependent on the noise level of the images (Vishnevskiy et al., 2016). All other registration hyperparameters were kept constant across the different (dis)similarity measures



(a) (b) (c)

Figure 4.3: (a - c) Three of the ten phases used in the Thoracic 4D CT experiment



(a) (b) (c)

(d) (e)

Figure 4.4: (a) CT image, (b) MR-PD image, (c) MR-T1 image, (d) MR-T2 image and (e) PET image used in the RIRE experiment.

4.2.4 RIRE

We performed multimodal experiments using the RIRE dataset which contains 18 patients with between three and five modalities (CT, PET, MR-T1, MR-T2

and MR-PD, Fig. 4.4) (West et al., 1997). Fiducial markers and a stereotactic frame allowed to determine anatomical landmarks which were employed to calculate the gTRE in our experiments.

To allow for a better convergence two rigid transformations were applied. First, a transformation with three translational degrees of freedom followed by a transformation with three additional rotational degrees of freedom. Given the multimodality of the data we compared one pairwise similarity measure ($\mathcal{S}_{MI}$) and three groupwise measures ($\mathcal{S}_{AAMI}$, $\mathcal{S}_{GAMI}$ and $\mathcal{S}_{HAMI}$). Since it is clear that $\mathcal{S}_{AAMI}$ will fail to obtain a representative template image given the range differences present in the data, we chose to normalize all images in our experiments between 0 and 1 prior to registration.

4.3 Results and discussion

4.3.1 Thoracic 4D CT

The results for the experiments on the 4D-CT data can be found in Table 4.1. It can be seen that there is no difference in the results for $\mathcal{S}_{AAMI}$, $\mathcal{S}_{GAMI}$ and $\mathcal{S}_{HAMI}$. We postulate that this is due to the three Pythagorean averages returning the same value when all elements in the set over which is averaged are identical. For monomodal data in perfect alignment, the three averages should return the same image and thus the convergence near the global minimum becomes the same. It can be seen that entropy based similarity measures ($\mathcal{S}_{MI}$, $\mathcal{S}_{AAMI}$, $\mathcal{S}_{GAMI}$ and $\mathcal{S}_{HAMI}$) performed best. Groupwise registrations took approximately twice as long to run compared to pairwise registrations.

Table 4.1: Results for the registration of the 4D CT of the lungs. The values, expressed in mm, correspond to the groupwise target registration error (gTRE)

Method	Mean ± stdev	Median/max
MSD	1.25 ± 0.29	1.18/1.80
MI	1.21 ± 0.25	1.17/1.60
SV	1.32 ± 0.40	1.22/2.43
AAMI	1.21 ± 0.25	1.18/1.61
GAMI	1.21 ± 0.25	1.18/1.61
HAMI	1.21 ± 0.24	1.18/1.61

Table 4.2: Results for the registration of the RIRE dataset. The values in the second and third column, expressed in mm, correspond to the groupwise target registration error (gTRE). In the fourth column the number of misregistrations are reported

Method	Mean ± stdev	Median/max	Misregs
MI	2.34 ± 0.81	2.18/4.30	0
AAMI	3.47 ± 1.85	2.85/7.50	2
GAMI	2.47 ± 0.61	2.42/3.55	0
HAMI	2.69 ± 1.27	2.65/6.85	0

4.3.2 RIRE

The results of the experiments on the RIRE dataset can be found in Table 4.2. The registration was considered to be a misregistration when the gTRE of a patient was more than 8mm, which corresponds to the largest voxel spacing of the images under study following Tomaževič et al. (2012). This occurred for $\mathcal{S}_{AAMI}$ in patient 008 and patient 105. These patients were excluded in the statistical analysis for all measures to allow for a fair comparison.

In addition to leading to two misregistrations when employing $\mathcal{S}_{AAMI}$, the gTRE is worse. The results seem to confirm that arithmetic averaging (even after normalizing the data) is not suitable for multimodal registration. Although the results for $\mathcal{S}_{HAMI}$ appear to be slightly worse, this is mainly caused by a single outlier, and needs to be investigated further.

4.4 Conclusion

In this work we presented two novel groupwise similarity measures for multimodal data based on the geometric and harmonic average as a template image. Experiments on a monomodal dataset containing intrasubject thoracic 4D-CT acquisitions showed no difference compared to arithmetic averaging, while experiments on a multimodal dataset containing intrasubject acquisitions of the brain showed both improved accuracy and robustness. We concluded that arithmetic averaging is not suitable for multimodal data in groupwise registrations. The limitation of current work is its application on a limited number of datasets and modalities. Further work is needed to validate the proposed methodology

on other modalities.

CHAPTER 5

Intrasubject multimodal groupwise registration with the conditional template entropy

Abstract

Image registration is an important task in medical image analysis. Whereas most methods are designed for the registration of two images (pairwise registration), there is an increasing interest in simultaneously aligning more than two images using *groupwise* registration. Multimodal registration in a groupwise setting remains difficult, due to the lack of generally applicable (dis)similarity measure. In this work, a novel similarity measure for such groupwise registration problems is proposed. The measure calculates the average of the conditional entropy between each image in the group and a representative template image constructed iteratively using principal component analysis. The proposed measure is validated in extensive experiments on synthetic and intrasubject clinical image data, which included monomodal data of the lungs (CT) and multimodal data of the brain (CT-MR-PET) and head and neck (MR, MR-CT). These experiments showed equivalent or improved registration accuracy compared to other state-of-the-art (dis)similarity measures and improved transformation consistency compared to pairwise mutual information.

Based upon: Polfliet, M., Klein, S., Huizinga, W., Paulides, M. M., Niessen, W. J., & Vandemeulebroucke, J. (2018). Intrasubject multimodal groupwise registration with the conditional template entropy. *Medical image analysis*, 46, 15-25.

5.1 Introduction

Biomedical image registration is the process of spatially aligning medical images, allowing for an accurate and quantitative comparison. An increasing number of image analysis tasks calls for the alignment of multiple (more than two) images. Examples include the joint analysis of tissue properties using multi-parametric MRI (Huizinga et al., 2016; Wells et al., 2015), spatio-temporal motion estimation from dynamic sequences (Metz et al., 2011; Vandemeulebroucke et al., 2011), atlas construction (Fletcher et al., 2009; Joshi et al., 2004; Wu et al., 2011) and population analyses (Geng et al., 2009).

One approach to perform such a registration task would be to take one image in the group as a reference and register all other images to this reference in a *repeated* pairwise manner. However, such an approach has two distinct shortcomings. First, the choice of the reference image inherently biases the resulting transformations and subsequent data analysis towards the chosen reference. Secondly, only a fraction of the total information available within the group of images is used in each pairwise registration, possibly leading to suboptimal results.

An alternative is to perform a *groupwise* registration in which all transformations are optimized simultaneously. Transformations are expressed with respect to a common reference space, thereby removing the need for choosing a particular reference image, and the bias associated with that choice. Additionally, a global cost function simultaneously takes into account all information in the group of images. In this work we will address such groupwise similarity measures for multimodal registration problems.

Multimodal intensity-based pairwise registration is commonly solved using mutual information (MI) (Collignon et al., 1995; Viola and Wells III, 1995; Wells et al., 1996), since it assumes a stochastic relationship between the two images to be registered. Extending MI to groupwise registration leads to a high-dimensional joint probability density function with an exponentially increasing number of histogram bins. Sparsity becomes a major concern as the number of images grows larger and limits the application to small groups of images (Wachinger and Navab, 2013).

A number of alternatives have been proposed to perform multimodal groupwise registration. Orchard and Mann (2010) proposed to use a Gaussian mixture model instead of histograms to approximate the joint probability density functions and Spiclin et al. (2012) approximated the joint probability density functions with a nonparametric approach based on a hierarchical intensity-space subdivision scheme. However, both approaches remain limited by the sparsity

in the joint intensity space and perform poorly for large groups of images.

Alternatively, one could represent the intensities as a graph and relate the length of such a graph to the entropy of the images (Hero et al., 2002). Such an approach requires a computationally expensive optimization for the construction of the graph and is not continuously differentiable, making gradient-based optimization difficult.

Zöllei et al. (2005) proposed the use of a voxelwise stack entropy. Herein, the intensities of all separate images in the group at a given sampled coordinate are grouped into a one-dimensional probability density distribution. For each sampled coordinate, the entropy is calculated and summed. However, for a low number of images in the group, the probability density functions are sparse which limits its use to larger groups of images.

Wachinger et al. (2007) proposed to accumulate all pairwise estimates of mutual information for all possible pairs of images in the group under consideration. Such an approach leads to a computation time which is proportional to the square of the number of images, making its application to larger groups of images increasingly difficult.

Joshi et al. (2004) developed an interesting dissimiliarity measure where the mean squared differences is used as a pairwise dissimiliarity measure to compare every image in the group to the voxel-wise average image. Herein the average image is updated in each iteration. They applied the method to monomodal brain atlas construction and it has also been applied to thoracic 4D CT data (Metz et al., 2011) and 4D ultrasound of the liver (Vijayan et al., 2014). The approach carries a number of advantages, such as the linear scaling of the computational complexity with respect to the number of images in the group and the possibility to parallelize the algorithm, making it feasible for both small and large groups of images. Bhatia et al. (2007) proposed to use the normalized mutual information (Studholme et al., 1999) as a pairwise similarity measure and the average image as a template image on monomodal intersubject data. The measure was termed the average normalized mutual information and has been used (together with the average mutual information) in subsequent literature as a similarity measure for multimodal groupwise registrations (Ceranka et al., 2018; Hallack et al., 2014; Huizinga et al., 2016; Polfliet et al., 2016, 2017). However, the use of the average image as the template image might not be appropriate in multimodal data with intensities of varying scales, ranges and contrast.

In this work a novel similarity measure, the conditional template entropy (CTE), is introduced for multimodal groupwise registration based on this principle of pairwise similarity with respect to a template image. Following the

original formulation by Joshi et al. (2004), we first design a suitable pairwise measure to be used in the comparison of the template image and every image in the group. Afterwards we investigate the use of a template image based on principal component analysis.

Given the linear scaling of the computational complexity, the proposed similarity measure can be applied to a wide range of intrasubject multimodal groupwise registration problems, for both small and large groups of images, and can be used as a general purpose measure. The proposed measure is validated in extensive experiments on synthetic and intrasubject clinical data, demonstrating equivalent or improved registration accuracy compared to other state-of-the-art methods and improved transformation consistency compared to pairwise MI.

5.2 Materials and methods

5.2.1 Pairwise registration

In pairwise registration, a target (moving, floating) image is registered to a reference (fixed, source) image. The transformation $\mathcal{T}_{\boldsymbol{\theta}} : \Omega_R \subset \mathbb{R}^d \rightarrow \Omega_T \subset \mathbb{R}^d$, parameterized by $\boldsymbol{\theta}$, needs to be determined that maps coordinates from the reference image domain, Ω_R, to the target image domain, Ω_T (Fig. 5.1(a)). The registration can be defined as an optimization (minimization) problem

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \mathcal{C}\left(I_R, I_T; \mathcal{T}_{\boldsymbol{\theta}}\right) \quad . \tag{5.1}$$

Here, are $I_T : \Omega_T \subset \mathbb{R}^d \rightarrow \mathbb{R}$ and $I_R : \Omega_R \subset \mathbb{R}^d \rightarrow \mathbb{R}$ the intensity functions of the target and reference image. $\mathcal{C}$ is the cost function or objective value of the registration problem, which is often represented as a weighted sum of a (dis)similarity measure, $\mathcal{S}$ or $\mathcal{D}$, and a regularization term, $\mathcal{R}$, such that

$$\mathcal{C} = \mathcal{D} + \lambda \mathcal{R} \quad , \tag{5.2}$$

in which λ is the weight for the regularization.

5.2.2 Mutual information

In the pairwise approach, mutual information (MI) (Collignon et al., 1995; Viola and Wells III, 1995; Wells et al., 1996) is defined as

$$\mathcal{S}_{MI}\left(I_R, I_T; \mathcal{T}_{\boldsymbol{\theta}}\right) = H\left(I_R\right) + H\left(I_T \circ \mathcal{T}_{\boldsymbol{\theta}}\right) - H\left(I_R, I_T \circ \mathcal{T}_{\boldsymbol{\theta}}\right) \quad . \tag{5.3}$$

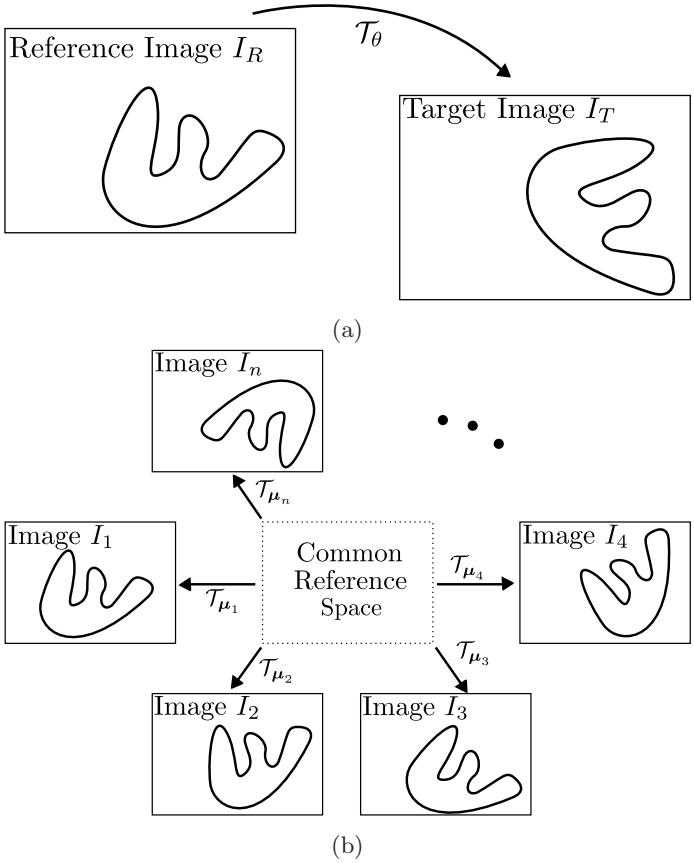


Figure 5.1: Graphical illustration for (a) a pairwise registration and (b) a groupwise registration.

Here, $H(\cdot)$ and $H(\cdot,\cdot)$ refer to, respectively, the marginal and joint entropy of the marginal and joint intensity distributions, often calculated via normalized histograms. In Eq. (5.3), the first term expresses the complexity of the reference image and the second term is the entropy of the target image mapped onto the reference, which favors transformations that map onto complex parts of the target image. The final term expresses the complexity of the shared or common relationship between the reference and target image. It is maximized when the (statistical or stochastic) relationship is stronger and thus less complex (Wells et al., 1996).

Following Maes et al. (1997), MI can be rewritten in terms of the conditional entropy (CE)

$$\mathcal{S}_{MI}\left(I_R, I_T \circ \mathcal{T}_{\boldsymbol{\theta}}\right) = H\left(I_R\right) - H\left(I_R | I_T \circ \mathcal{T}_{\boldsymbol{\theta}}\right) \quad . \tag{5.4}$$

The conditional entropy $H\left(A|B\right)$ describes the amount of information that remains in a random variable A once the random variable B is known. With the entropy of the reference image being independent of the transformation parameters, maximization of the negated conditional entropy and maximization of the mutual information lead to equivalent solutions of the registration problem.

5.2.3 Groupwise registration

In groupwise registration we consider a group of n images for which the transformations to a common reference frame are unknown. We can consider the following optimization (minimization) problem to determine these transformations:

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{C}\left(I_1, \ldots, I_n; \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, \mathcal{T}_{\boldsymbol{\mu}_n}\right) \quad , \tag{5.5}$$

where $\mathcal{T}_{\boldsymbol{\mu}_i} : \Omega \subset \mathbb{R}^d \rightarrow \Omega_i \subset \mathbb{R}^d$ the transformation that maps the coordinates from the common reference domain, Ω, to the domain of the ith image, Ω_i, and $\boldsymbol{\mu}_i$ are the parameters that define $\mathcal{T}_{\boldsymbol{\mu}_i}$ (Fig. 5.1(b)). For the optimization, $\boldsymbol{\mu}$ is defined as a vector formed by the concatenation of all separate transformation parameters, $\boldsymbol{\mu}_i$. $I_i : \Omega_i \subset \mathbb{R}^d \rightarrow \mathbb{R}$ is the continuous intensity function associated to the ith image.

5.2.4 Template construction

Joshi et al. (2004) proposed the following formulation for monomodal groupwise registration, in which both the transformation parameters and a template image

are optimized

$$\hat{\boldsymbol{\mu}}, \hat{J} = \arg\min_{\boldsymbol{\mu}, J} \frac{1}{n|S|} \sum_{i=1}^{n} \sum_{\boldsymbol{x} \in S} \left(I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i} (\boldsymbol{x}) - J(\boldsymbol{x}) \right)^2 \quad , \tag{5.6}$$

with J the continuous intensity function of a template image, $\boldsymbol{x}$ the coordinate samples drawn from the image domain and S the set of these samples. The template image can be interpreted as being the image that is most similar to the other images in the group in terms of the mean squared differences. For a given value of the transform parameters, the optimization with respect to the template image J was solved analytically to be the average image

$$J(\boldsymbol{x}) = \overline{I}_{\boldsymbol{\mu}} (\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i} (\boldsymbol{x}) \quad . \tag{5.7}$$

As such, the registration problem in Joshi et al. (2004) is reduced to

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \frac{1}{n|S|} \sum_{i=1}^{n} \sum_{\boldsymbol{x} \in S} \left(I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i} (\boldsymbol{x}) - \overline{I}_{\boldsymbol{\mu}} (\boldsymbol{x}) \right)^2 \quad . \tag{5.8}$$

5.2.5 The conditional template entropy

In this work, a novel similarity measure for multimodal groupwise registration is proposed, based on this paradigm in which similarity of the group of images is measured with respect to an iteratively updated template image. Considering the interpretation of the entropy terms given in Section 5.2.2, we propose to measure similarity using the negated joint entropy of each image in the group with the template image, favoring transformations for which the template explains the group of images well; and the marginal entropies of each image in the group, encouraging transformations that map onto complex parts of the images in the group. Note that this is equivalent to a formulation based on the conditional entropy:

$$\begin{aligned} \hat{\boldsymbol{\mu}}, \hat{J} &= \arg\max_{\boldsymbol{\mu}, J} \frac{1}{n} \sum_{i=1}^{n} H\left(I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}\right) - H\left(J, I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}\right) \\ &= \arg\max_{\boldsymbol{\mu}, J} - \frac{1}{n} \sum_{i=1}^{n} H\left(J | I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}\right) \quad . \end{aligned} \tag{5.9}$$

Observing the resulting measure, one can notice the resemblance with a formulation based on mutual information. The difference lies in the absence of the marginal entropy of the template image, $H(J)$. As we will demonstrate, this term counteracts the alignment of the group of images. A representative template image is likely to grow sharper when converging towards the optimal registration solution, leading to a reduced complexity of its intensity distribution and a decrease in the marginal entropy, which is opposite of the desired optimization behavior. The proposed method based on conditional entropy as shown in Eq. (5.9) eliminates this problem.

To find the appropriate template image, we revisit Eq. (5.6) where the template image could be obtained analytically as the average image. Unfortunately, Eq. (5.9) cannot be solved analytically with respect to the template image, J, for a given set of transformations if the trivial solution of a constant template image with a single intensity is excluded. Hypothetically, one could set up an optimization scheme where the template image is predefined by a functional relationship and weights corresponding to the images in the group. Herein, the optimization of the transformation parameters could be alternated with the optimization of the weights for the template image. Such nested optimization is error-prone and costly, and undesirable in this context.

Alternatively, instead of maximizing Eq. (5.9), we propose a more pragmatic approach which maximizes the variance in the template image. By defining J as the linear combination of the images in the group, principal component analysis (PCA) can be used to find the weights associated to the images.

PCA defines a linear transformation from a given high-dimensional space to a low-dimensional subspace whilst retaining as much variance as possible. In this work, PCA is performed with each sampled coordinate as a separate observation and the different images in the group corresponding to different features. The transformation to the 1-dimensional subspace along which the most variance is observed, is given by the eigenvector associated with the largest eigenvalue. As such, the elements of this eigenvector can serve as the weights for the construction of the template image.

$$J\left(\boldsymbol{x}\right) = I_{\boldsymbol{\mu}}^{PCA}\left(\boldsymbol{x}\right) = \sum_{i=1}^{n} v_{i,\boldsymbol{\mu}}\ I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}\left(\boldsymbol{x}\right) \quad . \tag{5.10}$$

Here, $\boldsymbol{v}_{\boldsymbol{\mu}}$ is the eigenvector associated with the largest eigenvalue and the subscript $\boldsymbol{\mu}$ is added to show its dependence on the transformation parameters. This template image, based on the principal component of the PCA, will hereafter be referred to as the principal component image.

Maximizing the variance in the template image carries a number of advantages over simple averaging. First, no need exists to preprocess the images with different intensity ranges (to avoid a single image overpowering the template) as it is inherently robust to such range differences. Secondly, assuming a quasi-linear relationship exists between the images, a maximization of the variance leads to a maximization of the contrast-to-noise ratio. Consider the case when the group of images is misaligned. This leads to smaller weights associated to those images that are more misaligned (due to a low correlation with the other images) and, subsequently, reduced noise in the template image (Melbourne et al., 2007). Similarly, when the images in the group are close to or at alignment, images with high noise content (due to acquisition settings or artefacts) will contribute less to the template image, resulting in improved contrast-to-noise ratio in the template image compared to simple averaging. Finally, PCA is robust to regions of interest where inversion of contrast occurs. This takes place in contrast enhanced imaging where a tumour might appear hypodense initially and hyperdense after contrast agent uptake, or in longitudinal pediatric brain MR imaging due to cortical organization and myelination of the white matter. Here, simple averaging could reduce the contrast-to-noise ratio of the template image in these regions of contrast inversion. PCA, on the other hand, is able to assign negative weights to certain images in the group. As a result, regions where contrast inversion takes place could propagate into the template image with higher contrast.

Combining Eq. (5.9) and (5.10) leads to a novel similarity measure, the conditional template entropy (CTE), where similarity is expressed as the sum of the conditional entropy between every image in the group and the principal component image:

$$\mathcal{S}_{CTE}\left(I_1, \ldots, I_n; \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, \mathcal{T}_{\boldsymbol{\mu}_n}\right) = -\frac{1}{n}\sum_{i=1}^{n} H\left(I_{\boldsymbol{\mu}}^{PCA} | I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}\right) \quad . \tag{5.11}$$

5.2.6 Optimization

The proposed similarity measure was implemented as part of the software package `elastix` (Klein et al., 2010) and is publicly available. An adaptive stochastic gradient descent was employed to minimize the cost function (Klein et al., 2009b). As such, the negated form of Eq. (5.11) is used, to allow a minimization to take place. The derivative of the proposed measure with respect to $\boldsymbol{\mu}$ was determined following the approach of Thévenaz and Unser (2000) in which B-splines were used as a Parzen windowing function such that the joint probability

density functions p_i between the template image and the ith image in the group become

$$p_i\left(\iota,\kappa;\boldsymbol{\mu}\right) = \alpha\sum_{\boldsymbol{x}}\left[\beta^m\left(\frac{\iota}{\epsilon_{PCA}} - \frac{I_{\boldsymbol{\mu}}^{PCA}\left(\boldsymbol{x}\right)}{\epsilon_{PCA}}\right)\beta^m\left(\frac{\kappa}{\epsilon_i} - \frac{I_i\left(\mathcal{T}_{\boldsymbol{\mu}_i}\left(\boldsymbol{x}\right)\right)}{\epsilon_i}\right)\right] . \tag{5.12}$$

Here, α is a normalization factor to obtain a density function, ϵ is related to the width of the histogram bin and β^m is a B-spline function of the order of m. ι and κ are the discretized intensities corresponding to the template image and images in the group, respectively. With B-splines fulfilling the partition of unity constraint (Thévenaz and Unser, 2000), we have

$$\sum_{\iota\in L_{PCA}}\sum_{\kappa\in L_i}\frac{\partial p_i\left(\iota,\kappa;\boldsymbol{\mu}\right)}{\partial\boldsymbol{\mu}} = 0 \quad \forall i \ , \tag{5.13}$$

where L_{PCA} and L_i are the discrete sets of intensities associated with the principal component and the ith image. This leads to

$$\frac{\partial\mathcal{S}_{CTE}}{\partial\boldsymbol{\mu}} = -\frac{1}{n}\sum_{i=1}^{n}\sum_{\iota\in L_{PCA}}\sum_{\kappa\in L_i}\frac{\partial p_i\left(\iota,\kappa;\boldsymbol{\mu}\right)}{\partial\boldsymbol{\mu}}\log\frac{p_i\left(\iota,\kappa;\boldsymbol{\mu}\right)}{p_{I_i}\left(\kappa;\boldsymbol{\mu}\right)} \tag{5.14}$$

With $p_{I_i}\left(\kappa;\boldsymbol{\mu}_i\right)$ the probability density function of the ith image. In Appendix A the derivative of the principal component image with respect to the transformation parameters is given.

5.2.7 Transformation degeneracy

Given the degeneracy of estimating n transformations for n images with an arbitrary global transformation, we chose to constrain our transformation following Bhatia et al. (2004) with

$$\frac{1}{n}\sum_{i=1}^{n}\mathcal{T}_{\boldsymbol{\mu}_i}(\boldsymbol{x}) = \boldsymbol{x}, \quad \forall\boldsymbol{x} \ , \tag{5.15}$$

i.e the sum of all transformations is the identity, effectively registering the group of images to the space for which the sum of the deformations at every coordinate is equal to zero. With Rosen's Gradient Projection Method (Luenberger, 1973) this is solved by setting

$$\frac{\partial\mathcal{C}}{\partial\boldsymbol{\mu}_i}' = \frac{\partial\mathcal{C}}{\partial\boldsymbol{\mu}_i} - \frac{1}{n}\sum_{j=1}^{n}\frac{\partial\mathcal{C}}{\partial\boldsymbol{\mu}_j} \ . \tag{5.16}$$

and using this projected gradient in the stochastic gradient descent optimization.

5.2.8 Regularization

Following Geng et al. (2009) we used a groupwise regularization term, the groupwise bending energy (GBE)

$$\mathcal{R}_{GBE}\left(\mathcal{T}_{\boldsymbol{\mu}_1},\ldots,\mathcal{T}_{\boldsymbol{\mu}_n}\right)=\frac{1}{|S|}\sum_{\boldsymbol{x}\in S}\frac{1}{n}\sum_{i=1}^{n}\sum_{l,m=1}^{d}\left|\left|\frac{\partial^2\mathcal{T}_{\boldsymbol{\mu}_i}(\boldsymbol{x})}{\partial x_l\partial x_m}\right|\right|^2. \tag{5.17}$$

Herein, d is the spatial dimension of the images. Regularization was performed in all clinical experiments with a deformable transformation model.

5.3 Data and experiments

A total of six experiments were conducted with two on synthetic data and four on clinical intrasubject data. Herein, the proposed conditional template entropy ($\mathcal{S}_{CTE}$) was compared to the average mutual information ($\mathcal{S}_{AMI}$)

$$\begin{aligned}&\mathcal{S}_{AMI}\left(I_1,\ldots,I_n;\mathcal{T}_{\boldsymbol{\mu}_1},\ldots,\mathcal{T}_{\boldsymbol{\mu}_n}\right)\\&=\frac{1}{n}\sum_{i=1}^{n}\left[H\left(\overline{I}_{\boldsymbol{\mu}}\right)+H\left(I_i\circ\mathcal{T}_{\boldsymbol{\mu}_i}\right)-H\left(\overline{I}_{\boldsymbol{\mu}},I_i\circ\mathcal{T}_{\boldsymbol{\mu}_i}\right)\right]\ .\end{aligned} \tag{5.18}$$

Furthermore, two auxiliary similarity measures were implemented to investigate complementary advantages of the proposed methodology, respectively the advantage of using the conditional entropy (S_{CE})

$$\begin{aligned}&\mathcal{S}_{CE}\left(I_1,\ldots,I_n;\mathcal{T}_{\boldsymbol{\mu}_1},\ldots,\mathcal{T}_{\boldsymbol{\mu}_n}\right)\\&=-\frac{1}{n}\sum_{i=1}^{n}H\left(\overline{I}_{\boldsymbol{\mu}}|I_i\circ\mathcal{T}_{\boldsymbol{\mu}_i}\right),\end{aligned} \tag{5.19}$$

and the advantage of using the principal component image (S_{PC})

$$\begin{aligned}&\mathcal{S}_{PC}\left(I_1,\ldots,I_n;\mathcal{T}_{\boldsymbol{\mu}_1},\ldots,\mathcal{T}_{\boldsymbol{\mu}_n}\right)\\&=\frac{1}{n}\sum_{i=1}^{n}\left[H\left(I_{\boldsymbol{\mu}}^{PCA}\right)+H\left(I_i\circ\mathcal{T}_{\boldsymbol{\mu}_i}\right)-H\left(I_{\boldsymbol{\mu}}^{PCA},I_i\circ\mathcal{T}_{\boldsymbol{\mu}_i}\right)\right]\ .\end{aligned} \tag{5.20}$$

For the clinical data, the four previously discussed groupwise similarity measures were used in addition to the PCA2 measure proposed in Huizinga et al. (2016) and pairwise MI (Eq. 5.3) as a baseline for comparison. PCA2 was proposed for the registration of images for which the intensity distribution could be represented into a low-dimensional subspace and is given as

$$\mathcal{D}_{PCA2}\left(I_1, \ldots, I_n; \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, \mathcal{T}_{\boldsymbol{\mu}_n}\right) = \sum_{i=1}^{n} i\lambda_i \ . \tag{5.21}$$

Herein, λ_i refers to the ith eigenvalue of the correlation matrix of the images in the group in decreasing order. In Huizinga et al. (2016) it was subsequently validated on monomodal and quantitative MRI image data for which such a low-dimensional subspace exists. PCA2 can be thus considered as a specialist dissimilarity measure specifically designed to register such images. To demonstrate the more generic nature of the proposed methodology, CTE was compared to PCA2 for both quantitative MRI and multimodal image data.

All registrations were performed in an intrasubject manner and the images were normalized by z-scoring to allow for a fair comparison to the similarity measures employing the average image. In the pairwise registration of a group of images, one image (the first in the sequence) was chosen as a reference to which all others were mapped. Note that other strategies for choosing the reference image in pairwise registrations for a group exist, such as the pre-contrast image in dynamic contrast enhanced sequences Kim et al. (2011), the end-expiration in 4D CT (Saito et al., 2009) or the mid-way image in computational anatomy (Reuter et al., 2010).

As the optimization strategy, interpolation algorithm, random sampler and transformation model is equivalent for all (dis)similarity measures, any difference in results can be solely attributed to the use of a different (dis)similarity measure.

The proposed methods were validated with two validation criteria. First, the groupwise target registration error (gTRE)

$$\mathrm{gTRE}\left(\boldsymbol{\mu}\right) = \frac{1}{n}\sum_{i\neq r}^{n}\frac{1}{|P_i|}\sum_{j}^{|P_i|}||\mathcal{T}_{i,r}\left(\boldsymbol{p}_{i,j}\right) - \boldsymbol{p}_{r,j}|| \tag{5.22}$$

was used as a measure for the accuracy of the registration with ground truth annotations of certain anatomical landmarks in the images. In Eq. (5.22), r is the index of the reference image, P_i the collection of landmarks in the ith image, $\mathcal{T}_{i,r}$ the transformation that maps the coordinates from the ith image to the reference

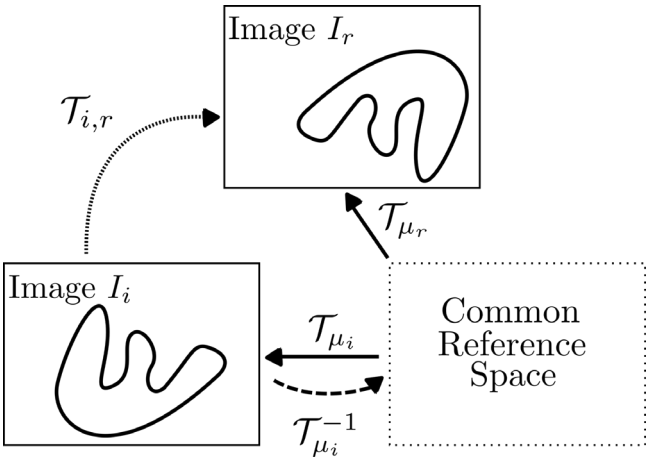


Figure 5.2: Composition of $\mathcal{T}_{\boldsymbol{\mu}_r}$ and $\mathcal{T}_{\boldsymbol{\mu}_i}^{-1}$ to obtain $\mathcal{T}_{i,r}$

image and $\boldsymbol{p}_{i,j}$ the jth landmark from the ith image. In a groupwise setting $\mathcal{T}_{i,r}$ was determined through the composition of the forward transformation, that maps the coordinates from the common reference space to the reference image, with the inverse transformation, that maps the coordinates from the ith image to the common reference space: $\mathcal{T}_{i,r} = \mathcal{T}_{\boldsymbol{\mu}_r} \circ \mathcal{T}_{\boldsymbol{\mu}_i}^{-1}$ (Fig. 5.2) (Metz et al., 2011). To allow for a fair comparison between pairwise and groupwise registrations, all validation measurements were performed in the same reference space, i.e. the same image which was chosen as a reference in the pairwise registrations.

Secondly, we computed the transitivity error (Christensen et al., 2006; Metz et al., 2011) to assess the quality of the transformation

$$\mathrm{Tra}\left(\boldsymbol{\mu}\right) = \frac{1}{|S|} \sum_{\boldsymbol{x} \in S} \sum_{i}^{n} \sum_{l \neq i}^{n} ||\mathcal{T}_{i,r}\left(\boldsymbol{x}\right) - \mathcal{T}_{i,l}\left(\mathcal{T}_{l,r}\left(\boldsymbol{x}\right)\right)|| \ . \tag{5.23}$$

The transitivity error measures the transitive property of the transformations in a group of images and can be interpreted as a measure for the consistency of the transformations in a groupwise setting. For *repeated* pairwise registrations the use of different reference images is required to measure the transitivity. The bias associated with the choice of a reference image will influence the results, whereas in groupwise registration, all transformations are estimated simultaneously and are inherently transitive (when the inverse transformation is available). As the

Table 5.1: The regularization weights used for each (dis)similarity measure and clinical dataset.

	Thoracic 4D CT	Carotid MR	Head&Neck	RIRE
PCA2	500	100	2×10^6	-
MI	0.02	50	100	-
AMI	0.05	100	2000	-
PC	0.2	100	2000	-
CE	0.01	100	5000	-
CTE	0.2	100	5000	-

inverse is approximated iteratively and the source for the transitivity error in the groupwise methods, no comparisons are made among the groupwise measures based on the transitivity error. The maximum transitivity error of the groupwise methods is reported and compared to the transitivity error of the pairwise method.

The cost function hyperparameters (the number of histogram bins and regularization weight) were chosen such that they optimized the mean gTRE per dataset. The different regularization weights are reported in Table 5.1. Due to the arbitrary sign of the projection vector for the principal component image, the number of histogram bins (used to calulate the entropy) are at least doubled compared to the number of histogram bins in registrations using the average image. Other optimization hyperparameters such as the spatial samples in the stochastic optimizer and the number of iterations were set to their default value. All registration hyperparameters in pairwise registrations were kept equal to those in the groupwise approach.

Results for the gTRE were compared in a pairwise manner among all (dis)-similarity measures (totaling 64 comparisons). The Wilcoxon signed-rank test was used for significance testing at a significance level of 0.05 adjusted by the Bonferroni correction for multiple comparisons.

5.3.1 Black&White

To investigate the effect the entropy term of the template image has on the optimization, an experiment was performed on synthetic data. Eleven identical black-and-white images were progressively and simultaneously translated along the horizontal axis and the (dis)similarity measures values were computed. A

Table 5.2: Summary of the registration parameters used in the experiments. Two values are reported for the number of histogram bins, separated by a forward slash. The first value reflects the number of bins used in pairwise registration and groupwise registrations based on the average image. The second value gives the number of bins used in groupwise registrations based on the principal component image. Values separated with a backward slash indicate multiple settings within the applied optimization strategy.

Dataset	Histogram bins	Resolutions	Grid spacing	Spatial samples	Iterations
Multimodal Cubes	32/96	2	6.0	2048	2000
Thoracic 4D CT	48/96	4	12.0	2048	2000\4000
Carotid MR	48/128	2	8.0	2048	2000
Head&Neck	64/144	2	64.0	2048	2000
RIRE	48/128	5\2	-	2048	2000

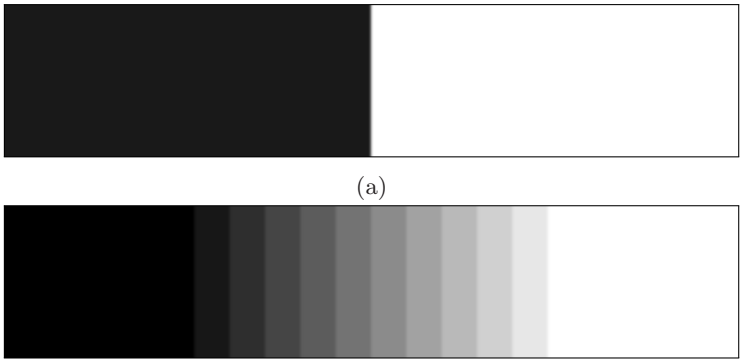
mask was used to keep the sampling domain constant. Fig. 5.3 shows a single black-and-white image and the average image of the group of images when they are at maximal displacement (15 mm).

5.3.2 Multimodal cubes

To further investigate registration accuracy, 100 registrations were performed on a group of six images (256 × 256 × 256 voxels) each containing two cubes, one surrounding the other. The intensities of the cubes and the backgrounds were set at random intensities to simulate a multimodal setting (Fig. 5.4). For each group of images a random set of deformable transformations was generated with a grid spacing of 8×8×8 voxels. The gTRE of the corners of the cubes was used to quantify the registration accuracy.

5.3.3 Thoracic 4D CT

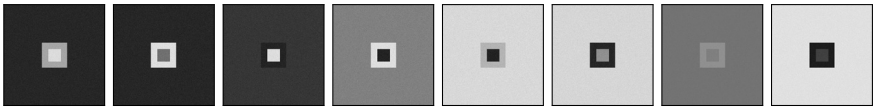
Thoracic 4D CT data (Fig. 5.5) was taken from the publicly available POPI and DIR-LAB datasets which include, respectively, 6 and 10 sequences of 10 respiratory phases each (Castillo et al., 2009; Vandemeulebroucke et al., 2011).



(a)

(b)

Figure 5.3: (a) A single black-and-white image (b) Average image of the group at their maximal misalignment.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 5.4: (a-f) A single slice of the six cubes used in the Multimodal Cubes experiment. (g) The average image and (h) the principal component image at alignment.

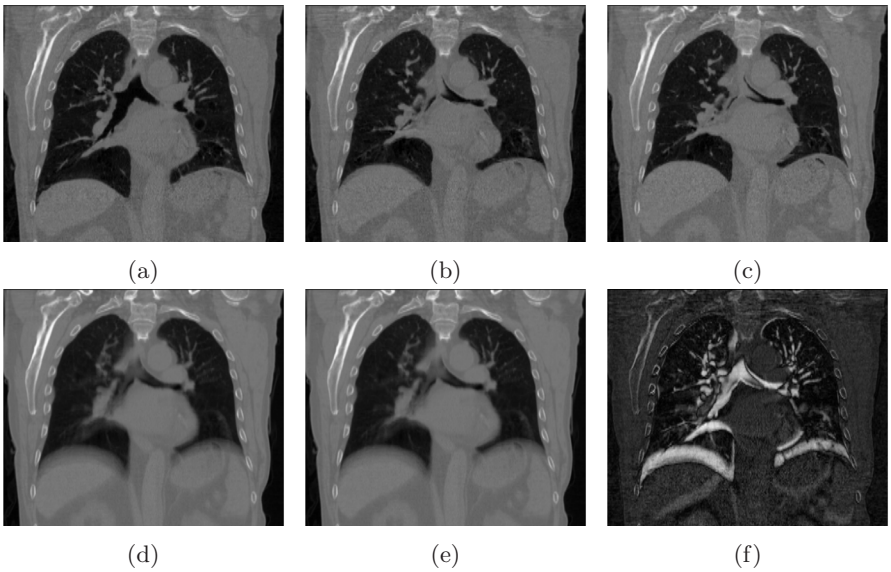


Figure 5.5: (a - c) Three of the ten phases used in the Thoracic 4D CT experiment. The images differ mainly in the position of the diaphragm and structures in the lungs due to breathing. (d) The average image at misalignment. (e) The principal component at misalignment. (f) Absolute difference image of the average and principal component image. Note that the largest differences occur in regions where motion is present (e.g. the diaphragm and trachea). The image contrast is optimized for the range of intensities present in each individual image.

Thoracic 4D CT data is often considered as monomodal data. However, minor intensity changes can occur due to changes in the voxel density in the lungs associated with the inhalation and exhalation of air (Sarrut et al., 2006) leading several authors to employ adapted or multimodal (dis)similarity measures for lung registration (Murphy et al., 2011b).

The POPI dataset contains three patients with 100 manually identified landmarks in the lungs for every breathing phase and three patients with 100 landmarks in end-inspiration and end-expiration phases with an inter-rater error of 0.5±0.9 mm. In the DIR-LAB dataset, all patients have 300 landmarks in the lungs for the inspiration and expiration phases and 75 in the four phases

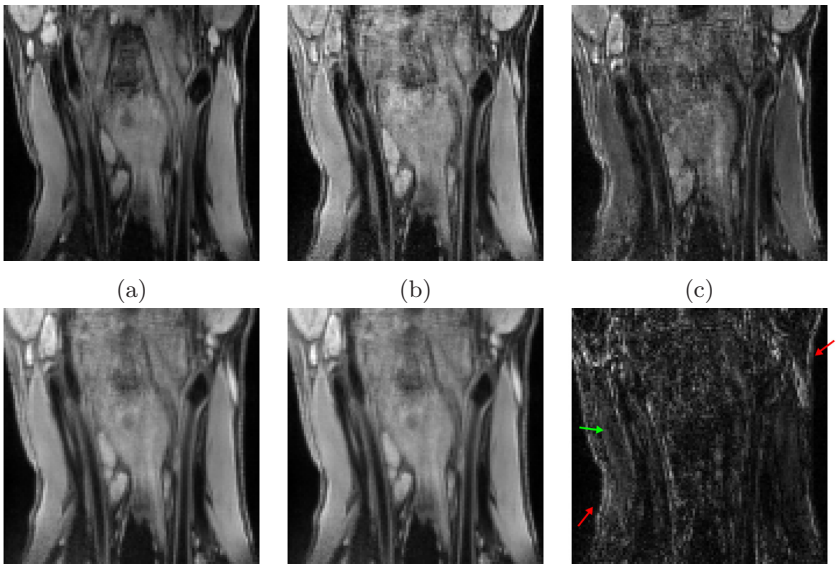
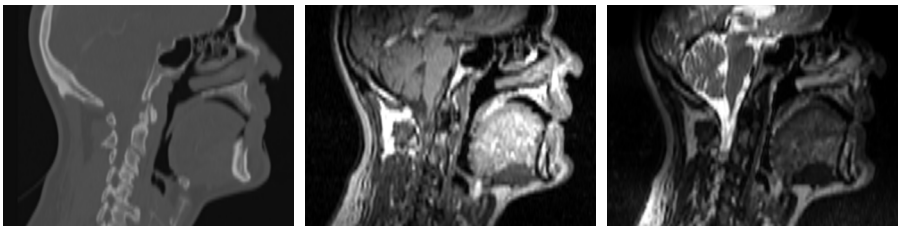


Figure 5.6: (a-c)Three of the five images used in the Carotid MR experiment. (d) The average image at misalignment. (e) The principal component at misalignment. (f) Absolute difference image of the average and principal component image. Note that the largest differences occur either at borders of structures due to motion, indicated by red arrows, or in homogeneous regions due to the multimodal nature of the data, indicated by a green arrow. The image contrast is optimized for the range of intensities present in each individual image.

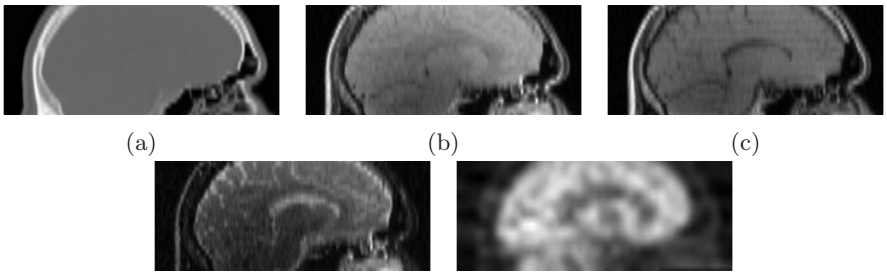
in between and an intra-rater error between 0.70 and 1.13 mm. Accuracy of the registration was determined using the gTRE with respect to the inspiration phase, the first image in the dynamic series.

A deformable registration was performed using cubic B-splines with a final grid spacing of 12.0 mm. Lung masks were used and obtained following (Vandemeulebroucke et al., 2012). For each resolution level 2000 iterations were performed, except for the last resolution where 4000 iterations were allowed.



(a) (b) (c)

Figure 5.7: (a) CT image, (b) MR-T1 image and (c) MR-T2 image used in the Head&Neck experiment.



(a) (b) (c)

(d) (e)

Figure 5.8: (a) CT image, (b) MR-PD image, (c) MR-T1 image, (d) MR-T2 image and (e) PET image used in the RIRE experiment.

5.3.4 Carotid MR

MR image sequences were acquired of the carotid artery by Coolen et al. (2015). The acquisitions were performed with a gradient echo MRI sequence for different flip angles and TE preparation times (Fig. 5.6). Each sequence consisted of five images and was performed for eight patients. The bifurcation of both carotid arteries was identified for each patient and consequently used as a landmark in the validation of the registration.

For this data we performed a deformable registration with cubic B-splines and a final grid spacing of 8.0 mm. van 't Klooster et al. (2013) has shown that a deformable registration is needed in such acquisitions of the carotid arteries. Masks around the carotid arteries were used as region of interest for registration.

5.3.5 Head&Neck

As part of radiotherapy planning, 22 patients underwent a CT, MR-T1 and MR-T2 imaging protocol of the head and neck region (Fortunati et al., 2014, 2015; Verhaart et al., 2014)(Fig. 5.7). In each acquisition between 15 to 21 landmarks were used to quantify the registration accuracy in terms of gTRE. The intra-rater variability of the landmarks was approximately 1mm.

Prior to registration, all images were resampled to the smallest voxel spacing present in the group of images. A deformable transformation was used in two resolution levels using cubic B-splines with a final grid spacing of 64.0 mm, as suggested by Fortunati et al. (2014).

5.3.6 RIRE

The RIRE dataset (West et al., 1997) includes 18 patients with up to five different imaging modalities of the brain (Fig. 5.8). All 18 patients had at least three of the following modalities available: CT, PET, MR-T1, MR-T2, MR-PD. Fiducial markers and a stereotactic frame were used to determine the ground truth transformations for CT to MR and PET to MR. Four to ten landmarks were available for each patient as a ground truth for the registrations and their target registration error was computed through the webform of the RIRE project, where rigid displacements between acquisitions were assumed.

To increase the robustness of the optimization, a two-step approach is used. First, a translation is optimized and used as an initialization for a second full rigid transformation with three translational and three rotational degrees of freedom. The registration was performed with five and two resolution levels, respectively. Similar to the Head&Neck dataset, preprocessing was performed by resampling the images in the group to the smallest voxel spacing.

The registration hyperparameters for the different experiments are summarized in Table 5.2.

5.4 Results

5.4.1 Synthetic data

The behavior of the similarity measures and its separate entropy components in the Black&White experiment are shown in Fig. 5.9 as a function of the translation. The Black&White experiment shows that the measure behavior of $\mathcal{S}_{AMI}$ and $\mathcal{S}_{PC}$ is equal to the behavior of the entropy of the images in the

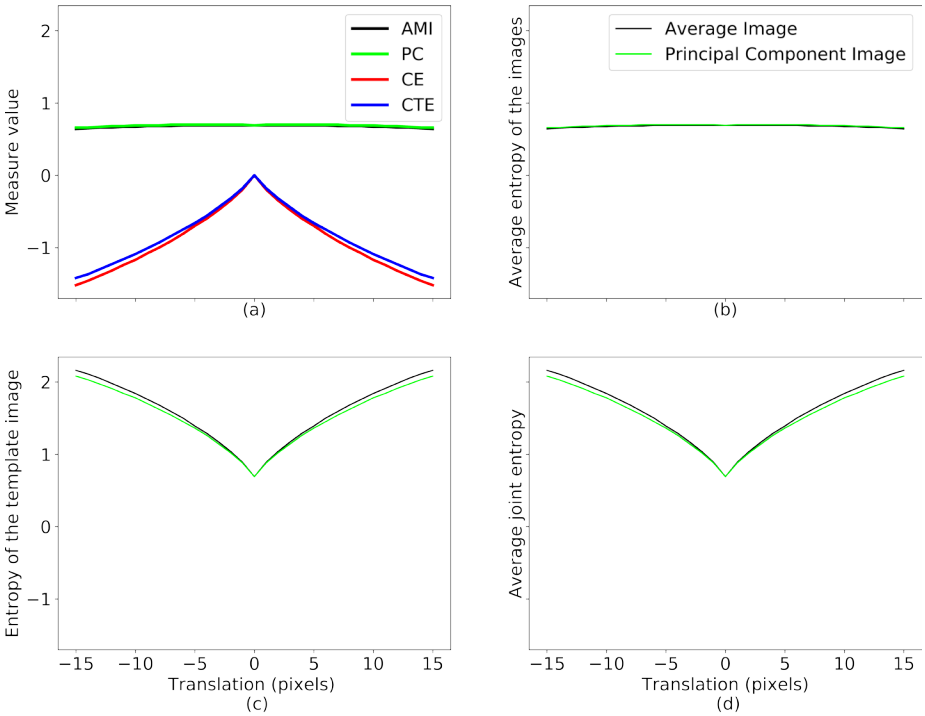


Figure 5.9: Results for the Black&White experiment where 11 black-and-white images were progressively and simultaneously translated. (a) The similarity measure values. (b) The average of the entropies of the images in the group. (c) The entropy of the template image. (d) The average of the joint entropies.

group. The contribution of the entropy of the template image completely cancels out the contribution of the joint entropy in $\mathcal{S}_{AMI}$ and $\mathcal{S}_{PC}$ as can be seen in Fig. 5.9(c-d). The resulting optimization is only driven by the complexity of the images in the group and not by their shared relationship.

The results for the Multimodal Cubes experiment are shown in Fig. 5.10. When comparing the similarity measures, $\mathcal{S}_{CTE}$ (1.71±0.11 mm) significantly outperformed all other entropy-based groupwise measures (2.80±0.32 mm, 2.73± 0.34 mm and 1.74±0.11 for $\mathcal{S}_{AMI}$, $\mathcal{S}_{PC}$ and $\mathcal{S}_{CE}$ respectively).

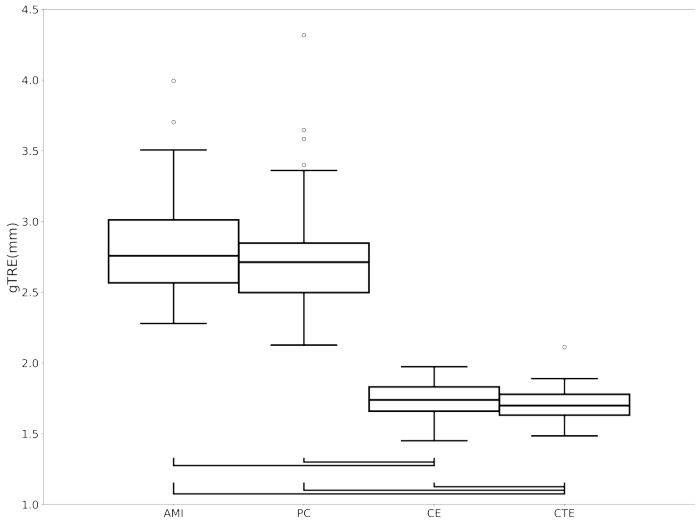


Figure 5.10: Boxplots for the results of the Multimodal Cubes experiment. Significant differences between two methods are indicated with black bars below the boxplots.

5.4.2 Clinical data

Results for the gTRE in experiments on clinical data are visualized with boxplots in Fig. 5.11&5.12.

For the experiments on the Thoracic 4D CT and Carotid MR datasets (Fig. 5.11), no statistically significant differences were observed in terms of gTRE for the investigated information-based measures. In Fig. 5.13, we visualized the deformation fields generated by the groupwise (dis)similarity measures for the thoracic 4D CT. Qualitatively, it is difficult no notice any difference between them.

In the Head&Neck experiment (Fig. 5.12) the best results are achieved by $\mathcal{S}_{CTE}$ with a gTRE of 2.74 ± 1.17mm performing significantly better compared to $\mathcal{S}_{AMI}$, $\mathcal{S}_{PC}$ and $\mathcal{D}_{PCA2}$.

Pairwise $\mathcal{S}_{MI}$ performed best in the RIRE experiment (Fig. 5.12) with a gTRE of 2.29 ± 0.72mm ($\mathcal{S}_{CTE}$, 2.33 ± 0.57mm), but no significant differences were found compared to the other entropy-based measures. $\mathcal{D}_{PCA2}$ performs worst, with the differences being statistically significant. A group of images was

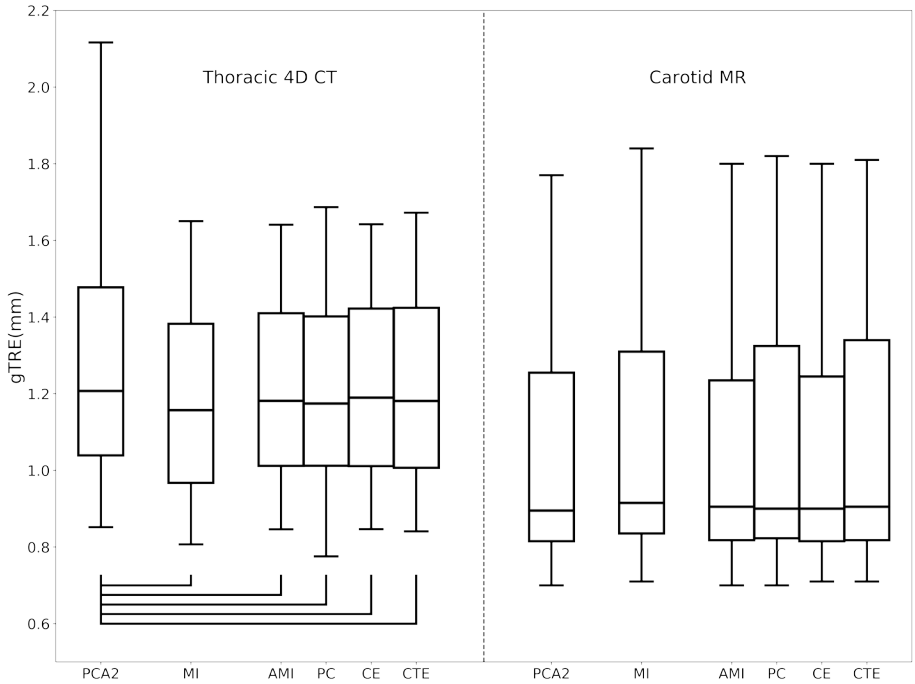


Figure 5.11: Boxplots for the results of the Thoracic 4DCT and Carotid MR experiment. Significant differences between two methods are indicated with black bars below the boxplots.

found to be misregistered following Tomaževič et al. (2012) when the gTRE is larger than the largest voxel spacing in the images. No misregistrations were obtained for $\mathcal{S}_{CTE}$, $\mathcal{S}_{CE}$ and $\mathcal{S}_{MI}$ whereas $\mathcal{S}_{AMI}$ and $\mathcal{S}_{PC}$ misregistered two patients and $\mathcal{D}_{PCA2}$ misregistered 14 patients.

In all four experiments on clinical data, pairwise MI performed worst in terms of transitivity, whereas the transitivity error for groupwise measures reduced to (close to) zero (Table 5.3).

In Table 5.4, the values are given for the average runtime of the experiments performed in this work. The use of the conditional entropy does not induce an extra computational burden, whereas the use of the principal component images does. This discrepancy originates from an additional loop over the sam-

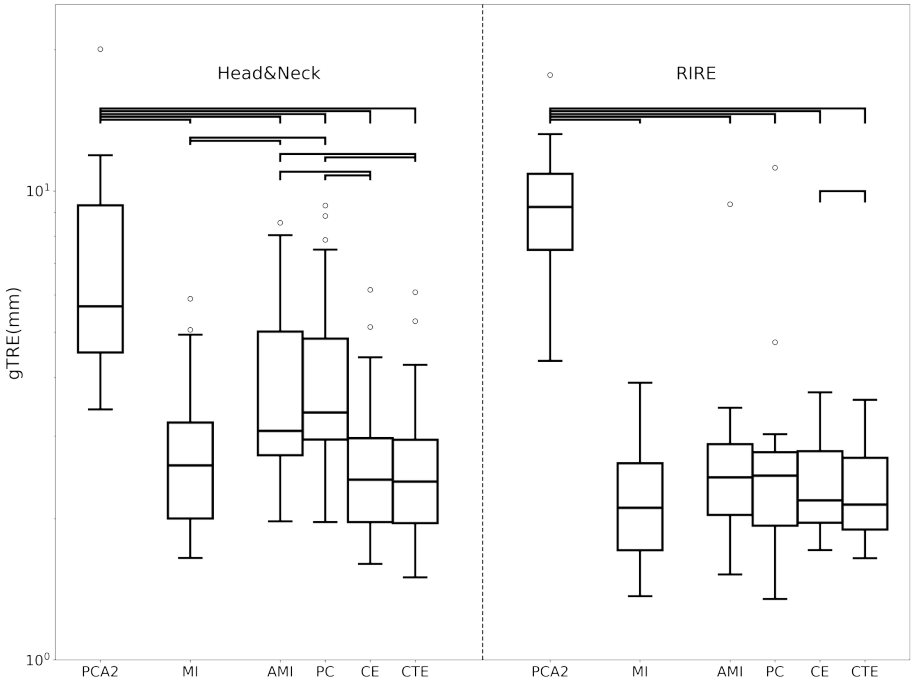
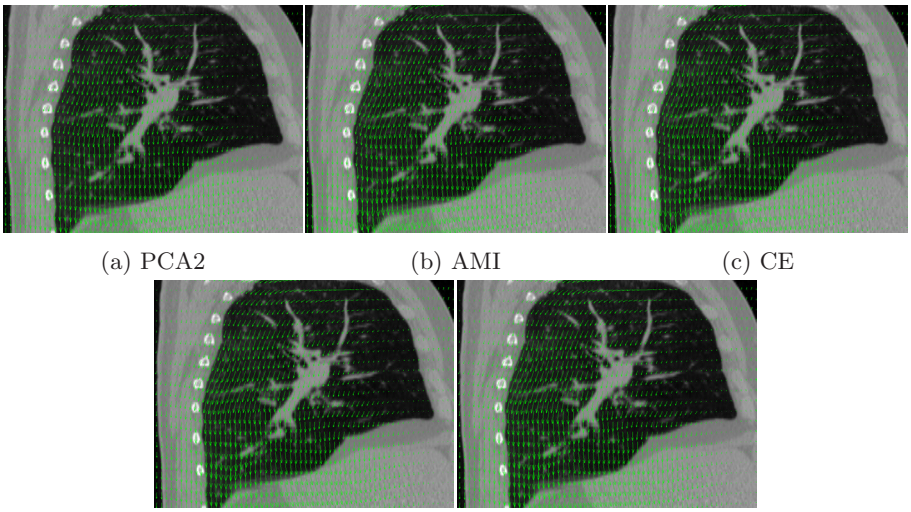


Figure 5.12: Boxplots for the results of the Head&Neck and RIRE experiment. Significant differences between two methods are indicated with black bars above the boxplots. Note the logarithmic scale on the y-axis.



(a) PCA2 (b) AMI (c) CE

(d) PC (e) CTE

Figure 5.13: Sagittal slice of the Thoracic 4D CT dataset overlayed with the resulting deformation field for the different measures under study, obtained from registrations where the field of view was limited to the lungs.

pled coordinates, needed to perform the PCA and determine the weights of the eigenvector. Note that for more complex registrations with a regularizer, the additional computation time is relatively small compared to the total cost.

5.5 Discussion

Results on the Thoracic 4D-CT and Carotid MR dataset showed equivalent performance of the proposed methodology compared to other state-of-the-art methods in terms of registration accuracy.

The results for the Multimodal Cubes, Head&Neck and RIRE results were consistent. In all three datasets the accuracy improved for the proposed formulation compared to $\mathcal{S}_{AMI}$, and the improvement was found to be statistically significant in the former two experiments. Throughout these experiments the behavior of the auxiliary (dis)similarity measures $\mathcal{S}_{CE}$ and $\mathcal{S}_{PC}$ was also consistent. Using the conditional entropy instead of mutual information led to a large

Table 5.3: Average transitivity errors for the clinical datasets. For the groupwise approaches, the maximum average transitivity error among all groupwise methods is reported. The values are given in mm.

	Thoracic 4D CT	Carotid MR	Head&Neck	RIRE
MI	5.65×10^{-1}	2.68×10^{-1}	2.14	1.47
Groupwise approaches	$< 3.39 \times 10^{-2}$	$< 7.66 \times 10^{-3}$	$< 1.85 \times 10^{-2}$	0

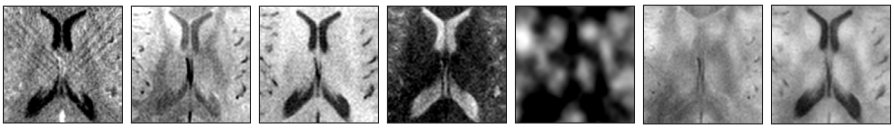
Table 5.4: Average runtime for the registrations in the different experiments. The values are given in minutes.

	Multimodal Cubes	Thoracic 4D CT	Carotid MR	Head&Neck	RIRE
PCA2	-	212	28	20	4
AMI	22	238	31	23	7
CE	22	252	31	23	7
PC	26	248	36	36	54
CTE	26	276	36	36	55

improvement, while using the principal component image improved the accuracy modestly. The combination of both contributions led to the best results in all three experiments compared to other groupwise (dis)similarity measures. As expected, the PCA2 measure performed poorly in multimodal registrations where a quantitative model or low-dimensional subspace does not exist.

In all experiments based on clinical data, the transitivity of the resulting transformations was compared to $\mathcal{S}_{MI}$ for groupwise approaches. These results emphasize the added value of the implicit reference space in multimodal groupwise registration. Whereas a pairwise approach has to perform two separate registrations with different reference images to obtain a concatenated transformation, in a groupwise approach all transformations are evaluated simultaneously and with a substantially lower transitivity error. These results are consistent with previous findings in monomodal data (Geng et al., 2009; Metz et al., 2011).

In summary, for experiments based on images where no or modest changes in intensity distributions are present ('Thoracic 4D-CT' and 'Carotid MR'), CTE



(a) (b) (c) (d) (e) (f) (g)

Figure 5.14: (a) CT image, (b) MR-PD image, (c) MR-T1 image, (d) MR-T2 image, (e) PET image, (f) average image and (g) principal component image when only the subregion of the ventricles is sampled for the RIRE experiment.

showed comparable performance to previously proposed groupwise methods and pairwise MI. In experiments with strongly varying intensity distributions ('Multimodal Cubes', 'Head&Neck' and 'RIRE'), CTE showed superior performance to previously proposed groupwise methods and performed on par to pairwise MI, with little to no transitivity error.

Fig. 5.5(f) and 5.6(f) highlight the differences in the average and principal component images. Herein, the absolute difference image between the average and principal component image is given in the 'Thoracic 4D CT' and 'Carotid MR' dataset, respectively, for a single patient. Herein, the largest differences occur in regions where the motion is greatest near moving structures or edges. This is consistent with previous work, where the principal component image was used to separate motion present in the images (Feng et al., 2016; Hamy et al., 2014; Melbourne et al., 2007). For multimodal registrations, the benefit of PCA over averaging can be seen by considering cases in which images with an inverted intensity profile are merged into the template image, as shown in Fig. 5.4(g-h) and Fig. 5.14. For the 'Multimodal Cubes' experiment, PCA lead to an increase of the contrast-to-noise ratio from 7.4 to 32.5 compared to simple averaging. Fig. 5.14 shows the average and principal component image when applied to the ventricles for an arbitrary patient in the RIRE dataset. With the T2 modality having an inverted intensity profile, the principal component image is able to retain the contrast in the template image. In the average image the intensities cancel out and the ventricles are poorly visible.

Two limitations should be stated with respect to current work. Firstly, only intrasubject data has been employed. Intersubject data is characterized by greater variability of intensity profiles and morphology, and has been reported to considerably increase the complexity of groupwise registration (Hamm et al., 2009; Tang et al., 2009). It remains to be verified how CTE would perform when confronted with such data.

Secondly, in this work a methodology was used where the images are deformed and compared to the template image in the implicit reference system. However, previous work has shown that deforming the template image to the images in the group suits a generative model better (Allassonnière et al., 2007; Ma et al., 2008). In methodologies where the template is deformed to the images in the group, no need exists to constrain the transformations to the average deformation space (Eq. 5.16). This was shown to be advantageous, as such constraints could exclude some legitimate results (Aganj et al., 2017). We expect the proposed measure to perform equally well in such frameworks as it is independent of the transformations that were used.

5.6 Conclusion

In this work we proposed a novel similarity measure for intrasubject multimodal groupwise registration, the conditional template entropy. The proposed measure was evaluated in experiments based on synthetic and clinical intrasubject data and showed equivalent or improved registration accuracy compared to other state-of-the-art (dis)similarity measures and improved transformation consistency compared to pairwise mutual information. These improvements were achieved mainly by the use of the conditional entropy, whereas the use of the principal component image contributed modestly in our experiments.

Acknowledgement

The research of S. Klein, W.J. Niessen and W. Huizinga is supported by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 601055, VPH-DARE@IT. Additionally, the authors would like to thank Dr. B.F. Coolen and Dr. A.J. Nederveen for providing the 'Carotid MR' data.

Appendix A: Derivative of principal component image

We determined the derivative of the principal component image with respect to the transformation parameters. The principal component image is given by

Eq. (5.10) and repeated here

$$I_{\boldsymbol{\mu}}^{PCA}(\boldsymbol{x}) = \sum_{i=1}^{n} v_{i,\boldsymbol{\mu}}\ I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}(\boldsymbol{x}) = \boldsymbol{v}_{\boldsymbol{\mu}}^T \boldsymbol{I}(\boldsymbol{x}) \quad . \tag{5.24}$$

Herein, $\boldsymbol{I}(\boldsymbol{x})$ is the column vector representing all image intensities across the group for a given sampled coordinate. The derivative becomes

$$\frac{\partial I_{\boldsymbol{\mu}}^{PCA}(\boldsymbol{x})}{\partial \boldsymbol{\mu}} = \frac{\partial \boldsymbol{v}_{\boldsymbol{\mu}}^T}{\partial \boldsymbol{\mu}} \boldsymbol{I}(\boldsymbol{x}) + \boldsymbol{v}_{\boldsymbol{\mu}}^T \frac{\partial \boldsymbol{I}(\boldsymbol{x})}{\partial \boldsymbol{\mu}} \quad , \tag{5.25}$$

Following de Leeuw (2007) for the derivative of an eigenvector:

$$\frac{\partial \boldsymbol{v}_{\boldsymbol{\mu}}}{\partial \boldsymbol{\mu}} = -\left(C - eI\right)^{+} \frac{\partial C}{\partial \boldsymbol{\mu}} v_{i,\boldsymbol{\mu}} \quad , \tag{5.26}$$

with C the correlation matrix of the intensities, similar to Huizinga et al. (2016), I the identity matrix, e the eigenvalue associated with $\boldsymbol{v}_{\boldsymbol{\mu}}$ and $^+$ the notation for the Moore-Penrose inverse (de Leeuw, 2007). The derivative of the correlation matrix is given as

$$\begin{aligned} \frac{\partial C}{\partial \boldsymbol{\mu}} = & \frac{1}{|S|-1} \Bigg(\frac{\partial \Sigma^{-1}}{\partial \boldsymbol{\mu}} \left(M - \overline{M}\right)^T \left(M - \overline{M}\right) \Sigma^{-1} \\ & + \Sigma^{-1} \frac{\partial M^T}{\partial \boldsymbol{\mu}} \left(M - \overline{M}\right) \Sigma^{-1} \\ & + \Sigma^{-1} \left(M - \overline{M}\right)^T \frac{\partial M}{\partial \boldsymbol{\mu}} \Sigma^{-1} \\ & + \Sigma^{-1} \left(M - \overline{M}\right)^T \left(M - \overline{M}\right) \frac{\partial \Sigma^{-1}}{\partial \boldsymbol{\mu}} \Bigg) \end{aligned} \tag{5.27}$$

Herein, M refers to the data matrix with the intensities of the images, $\overline{M}$ is the matrix with the average image intensity repeated along its columns, Σ is the diagonal matrix with the standard deviations of the images intensities as its diagonal elements. All notations correspond to those found in Huizinga et al. (2016) and we have ignored the derivative of the average image intensities likewise.

CHAPTER 6

Laplacian eigenmaps as a dissimilarity measure in multimodal groupwise registration

Abstract

In this work we propose a novel dissimilarity measure for multimodal groupwise registration based on Laplacian eigenmaps, a non-linear dimensionality reduction technique. Dissimilarity in the group of images is expressed as the magnitude of the Fiedler eigenvalue, the second generalized eigenvalue of the Laplacian of the graph constructed in the joint intensity space. The proposed methodology, combined with a stochastic gradient descent, allows for an efficient simultaneous registration of large groups of images. Furthermore, it can capture non-linear relationships, and alleviates issues often associated to entropy-based similarity measures due to bias fields in the images and sparsity in the joint intensity space. Experiments were performed on four datasets to validate the proposed methodology, one based on synthetic data, one on simulated data and two on clinical multimodal image data of the brain (MR) and knee (MR). Herein we show improved or equivalent performance compared to other state-of-the-art groupwise (dis)similarity measures in terms of registration accuracy, in addition to an efficient scaling of the computational complexity.

Based upon: Polfliet, M., Klein, S., Niessen, W. J., & Vandemeulebroucke, J. Laplacian Eigenmaps as a Dissimilarity Measure in Multimodal Groupwise Registration.

6.1 Introduction

Groupwise image registration has numerous applications, such as atlas construction (Fletcher et al., 2009; Joshi et al., 2004; Serag et al., 2012; Wu et al., 2011, 2016), multi-atlas based segmentation (Bhatia et al., 2007; Makropoulos et al., 2017), longitudinal population analysis (Davis et al., 2010; Huizinga et al., 2018) and spatiotemporal motion estimation (Metz et al., 2011; Royuela-del Val et al., 2016; Yigitsoy et al., 2011). However, registering groups of biomedical images with differences in the intensity distribution, due to anatomical changes or different acquisition protocols, remains challenging. Such multimodal groupwise registrations require suitable (dis)similarity measures which are capable of handling the intensity differences and are applicable to both small and large groups of images. Whereas pairwise multimodal registration is typically based on mutual information (Maes et al., 1997; Wells et al., 1996), its extension to groupwise registration, multivariate mutual information, is not straightforward. Using normalized histograms to estimate the entropy, leads to an exponentially increasing number of histogram bins. Herein, sparsity is of major concern and the images that can be registered using such an approach are limited to a small group.

A number of methods have been proposed to alleviate or circumvent the issue of sparsity in the joint intensity space. One can approximate the probability distribution differently using Gaussian clusters (Studholme and Cardenas, 2004) or Gaussian mixture models (Orchard and Mann, 2010). However, these approaches remain subject to sparsity in the joint intensity space for large groups. Alternatively, entropic graphs provide an estimation for the entropy without calculating the probability distribution (Neemuchwala et al., 2006). Herein, the entropy is estimated based on some length descriptor of a graph (minimal spanning tree length, k-nearest neighbors graph length, ...) in the joint intensity space. Other authors avoided the need to calculate high-dimensional joint entropies and compared every image in the group with every other image (Wachinger and Navab, 2013) or with some template image (Bhatia et al., 2007; Polfliet et al., 2016, 2018). Guyader et al. (2018) derived a similarity measure based on total correlation, a multivariate extension of mutual information, under the assumption that the image intensities are normally distributed.

Closely related to mutual information-based approaches is the framework of congealing or the stack entropy, where the sum of the voxelwise entropy is calculated (Learned-Miller, 2006; Zöllei et al., 2005). At each coordinate, the intensities among all images in the group are collected in a histogram and the histogram entropies at each sampled coordinate are summed. However, such an

approach has exclusively been applied to monomodal images (Balci et al., 2007; Warfield et al., 2002) and has not been validated for multimodal registration. Additionally, sparsity occurs when the number of images in the group is low.

In this work we propose a novel dissimilarity measure for multimodal groupwise registration in which we represent a joint intensity relationship in n-dimensional intensity space (where n is the number of images in the group) using manifold learning, a technique for non-linear dimensionality reduction. Dimensionality reduction attempts to reduce high-dimensional data in the observable space to low-dimensional data in the latent space while minimizing the loss of information. It is typically performed in the context of machine learning as a feature extraction step. Non-linear dimensionality reduction assumes that the data lies on or close to a low-dimensional manifold embedded in the high-dimensional space. It attempts to find the manifold onto which the high-dimensional data is best reduced. Several methodologies have been proposed to perform manifold learning, including local linear embedding (Roweis and Saul, 2000), isomap (Tenenbaum et al., 2000) and Laplacian eigenmaps (Belkin and Niyogi, 2003). We present a novel dissimilarity measure for groupwise multimodal registration based on the Laplacian eigenmaps algorithm, given its efficiency (Azampour et al., 2014) and similar interpretability to principal component analysis (Sharma et al., 2009). The measure is expressed as the magnitude of the second generalized eigenvalue from the Laplacian. Every sampled coordinate can be considered as an observation or sample in the joint intensity space and the dimension to be embedded is the group dimension, e.g. the cross-sectional dimension in population analysis, the temporal dimension in motion analysis or the multimodal dimension in multi-parametric MRI.

Dimensionality reduction algorithms have previously been applied to image registration and a short overview is provided. We highlight some key differences between the different algorithms and differentiate between two groups, those where the dimensionality reduction is applied to the spatial dimensions and those where it is applied to the group dimension.

First we will discuss *spatial* methods. In Wachinger and Navab (2012) and Azampour et al. (2014) a manifold of image patches was built using Laplacian eigenmaps where the dimensionality of the observable space was equal to the volumetric size of the image patches. An image patch with a size of three by three by three pixels would result in a 27-dimensional space to be reduced. The manifold was used to obtain a structural representation of multimodal images, allowing the resulting structural images to be registered with a monomodal (dis)similarity measure. Similarly, Guerrero et al. (2011) used Laplacian eigenmaps to localize landmarks based on image patches. The struc-

tural representation, obtained from the dimensionality reduction, could then be employed to predict the displacements. Aforementioned approaches, together with Baumgartner et al. (2014); Lombaert et al. (2014); Piella (2014); Zimmer et al. (2019), applied a dimensionality reduction algorithm where the dimensionality of the observable space was defined by the spatial size of the image or image patches.

For the *group* methods, Melbourne et al. (2007) and Polfliet et al. (2018) applied principal component analysis (PCA) to the group dimension to build template images, these images were then compared in a pairwise manner to the other images in the group with the cross-correlation and the conditional entropy, respectively. In all previously discussed methods, dimensionality reduction was performed as part of a multi-step approach where it typically was employed as a preprocessing to extract features or an improved intensity representation. Huizinga et al. (2016), on the other hand, proposed to use PCA, a *linear* dimensionality reduction technique, to register groups of quantitative MRI images. Herein, PCA was applied on the group dimension and it was assumed that more variance in the intensity correlation matrix can be explained by its first few eigenvalues when the images are closer to alignment. As a result, the eigenvalues themselves could be employed as an intensity-based (dis)similarity measure.

In this work we propose a novel dissimilarity measure for multimodal groupwise registration based on Laplacian eigenmaps, which does not require any pre- or postprocessing and we ensure its tractability through the use of a stochastic optimization approach. The proposed measure is evaluated in four datasets and compared to state-of-the-art multimodal groupwise (dis)similarity measures. Two datasets include synthetic or simulated data and two datasets include clinical intersubject data. This work extends our previous conference paper (Polfliet et al., 2017) presented at SPIE, Medical Imaging 2017. Compared to the conference paper, this work includes a method for automated hyperparameter estimation, experiments on new datasets, timing and smoothness results, and a hyperparameter robustness experiment. Additionally, we have investigated how different (dis)similarity measures behave as a function of the number of images.

6.2 Material and methods

6.2.1 Groupwise registration

In groupwise registration we consider a group of n images for which the transformations to a common reference frame are unknown. The following optimization problem can be considered to determine these transformations:

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{C}\left(I_1, \ldots, I_n; \mathcal{T}_{\boldsymbol{\mu}_1}, \ldots, \mathcal{T}_{\boldsymbol{\mu}_n}\right) \quad , \tag{6.1}$$

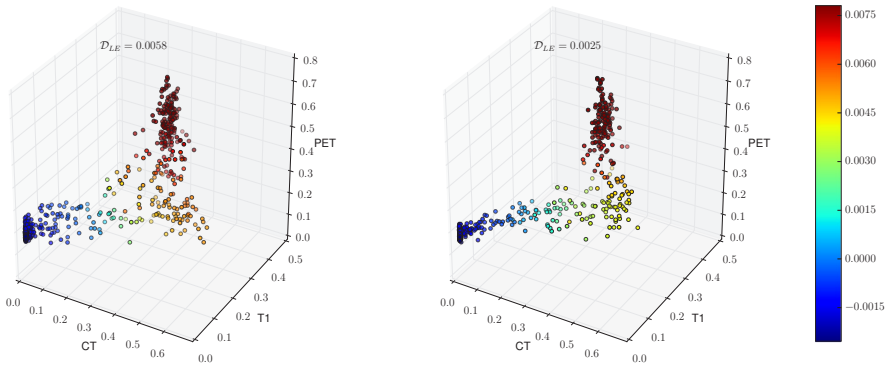
where $\mathcal{C}$ is the cost function, typically defined as the weighted sum of a dissimilarity measure, $\mathcal{D}$, and a regularizer, $\mathcal{R}$. The dissimilarity measure evaluates the misalignment of the transformed images, $I_i \circ \mathcal{T}_{\boldsymbol{\mu}_i}$. $\mathcal{T}_{\boldsymbol{\mu}_i} : \Omega \subset \mathbb{R}^3 \to \Omega_i \subset \mathbb{R}^3$ is the transformation that maps the coordinates from the common reference domain, Ω, to the domain of the ith image, Ω_i, and $\boldsymbol{\mu}_i$ are the parameters that define $\mathcal{T}_{\boldsymbol{\mu}_i}$. For the optimization, $\boldsymbol{\mu}$ is defined as a vector formed by the concatenation of all separate transformation parameters, $\boldsymbol{\mu}_i$. $I_i : \Omega_i \subset \mathbb{R}^3 \to \mathbb{R}$ is the continuous intensity function of the ith image assuming an appropriate interpolation scheme on the discrete image.

The cost function in Eq. (6.1) is typically calculated over a finite number, s, of coordinate samples, $\boldsymbol{x}_j \in \Omega$ and for each sample ($j = 1 \ldots s$) and image ($i = 1 \ldots n$) the intensity value $I_i\left(\mathcal{T}_{\boldsymbol{\mu}_i}\left(\boldsymbol{x}_j\right)\right)$ is computed. These intensity values can be represented by a single $s \times n$ matrix $\boldsymbol{M}$. We will denote the column vector $\boldsymbol{m}_i$ as the ith column and the row vector $\boldsymbol{m}^j$ as the jth row of the matrix $\boldsymbol{M}$.

6.2.2 Laplacian eigenmaps

Laplacian eigenmaps (LE), originally proposed by Belkin and Niyogi (2003), is a technique commonly employed to perform manifold learning and its procedure can be subdivided in three steps. In the context of image registration formulated in Eq. (6.1), these steps are given as:

1. Construct an undirected simple graph in which every row $\boldsymbol{m}^j$ of the $s \times n$ matrix $\boldsymbol{M}$ is an n-dimensional node (or vertex) in the graph spanning the joint intensity space. Each node is connected with an edge to its neighbours, determined by a nearest neighbour search.



(a) (b)

Figure 6.1: Joint intensity space with 2048 coordinate samples, s, for a group of images including a CT, MR-T1 and PET (a) before and (b) after alignment. The color coding represents the value of the first coordinate in the embedded manifold of the samples, i.e. $\boldsymbol{v}_1$. When the images are aligned, the spatial samples are less dispersed in the joint intensity space and they more closely conform to an intensity relationship which can be represented by, in this case, a one-dimensional manifold.

2. Construct the $s \times s$ weight matrix, $\boldsymbol{W}$, where

$$W_{ab} = \begin{cases} 0 & \text{if a, b are not connected} \\ e^{-\frac{||\boldsymbol{m}^a - \boldsymbol{m}^b||^2}{t}} & \text{if a, b are connected} \end{cases} \tag{6.2}$$

   Herein, t is a scalar hyperparameter of the procedure and referred to as the diffusion time. The diffusion time is discussed in further detail in Sec. 6.2.5

3. Embed the manifold into a lower dimension by optimally preserving the locality, i.e. the distance between the nodes, in the graph with respect to a suitable objective function.

Let v^j be such a mapping of $\boldsymbol{m}^j$ onto a 1-dimensional line. The third step in the algorithm is then achieved by minimizing the following objective function

$$\frac{1}{2} \sum_{a,b} \left(v^a - v^b\right)^2 W_{ab} = \boldsymbol{v}^T \boldsymbol{\mathcal{L}} \boldsymbol{v} \ , \tag{6.3}$$

where $\boldsymbol{\mathcal{L}} = \boldsymbol{D} - \boldsymbol{W}$ is the Laplacian of the graph and $\boldsymbol{D}$ the diagonal matrix of the graph with its elements $D_{aa} = \sum_b W_{ab}$. The minimization in Eq. (6.3) with respect to $\boldsymbol{v}$, under the constraint $\boldsymbol{v}^T \boldsymbol{D} \boldsymbol{v} = \mathbf{1}$ to remove the arbitrary scaling, can be reduced to a generalized eigenvalue problem

$$\boldsymbol{\mathcal{L}} \boldsymbol{v} = \lambda \boldsymbol{D} \boldsymbol{v} \ . \tag{6.4}$$

6.2.3 LE-based dissimilarity measure

We propose to use the Laplacian eigenmaps algorithm to model the (quasi-) manifold originating from the intensities of multimodal images in the joint intensity space. Specifically, the second smallest eigenvalue, λ_1, (the smallest eigenvalue is equal to zero) of the eigenvalue problem posed in Laplacian eigenmaps can be used as a dissimilarity measure which needs to be minimized,

$$\mathcal{D}_{LE} = \lambda_1 \ . \tag{6.5}$$

More commonly known as the Fiedler eigenvalue or the algebraic connectivity, the properties of this eigenvalue and associated eigenvector have been studied extensively in graph theory (Fiedler, 1973). Historically the magnitude of the algebraic connectivity was used to express the connectivity and robustness of the graph, where a lower value for the eigenvalue corresponds to a lower connectivity in the graph.

The proposed Laplacian eigenmaps-based measure is similar to the PCA-based measures from Huizinga et al. (2016). Indeed, Sharma et al. (2009) have previously shown that the eigenvalues obtained from LE are closely related to those of PCA. Where PCA maximizes the variance in the original coordinates, it can be shown that LE maximizes the variance in the embedded coordinates with the inverse of the eigenvalues from Eq. (6.4), i.e. λ^{-1}, expressing the explained variance.

We hypothesize that when the images are misaligned, the samples in the joint intensity space are dispersed more, and less variance in the embedded space is captured by the first eigenvalue of Eq. 6.4, leading to a larger value of the proposed dissimilarity measure in Eq. 6.5. This concept is illustrated in Fig. 6.1 where the joint intensity space for a group of three images is given, a CT image, an MR image and a PET image (taken from the RIRE database (West et al., 1997)). After alignment the intensities are less dispersed and more clustered. Note that while the interpretation is similar to the PCA-based measures, the proposed measure based on LE could capture more complex, non-linear relationships in the joint intensity space, potentially allowing for the alignment

of multimodal images. Moreover, the historical interpretation of the Fiedler eigenvalue to correspond to the connectivity and robustness of the graph can be understood from a clustering perspective. When the images are aligned, the nodes in the graph corresponding to different anatomical structures are tightly clustered in different regions of the joint intensity space. Here, the removal of a single node can break the graph apart in several subgraphs, making the graph fragile and its connectivity low. However, at misalignment, outlier nodes (corresponding to one anatomical structure in one image and to another structure in other images) are more likely to occur, which add additional connectivity between the intensity clusters.

The proposed dissimilarity measure based on LE can also be interpreted to be similar to the correlation ratio (CR). Roche et al. (1998a,b) proposed CR as an alternative to mutual information in multimodal pairwise registration. It was subsequently employed and extended to register US-MR image pairs (Rivaz et al., 2015; Roche et al., 2001), DT-MRI (Zvitia et al., 2010) and respiratory CT (Wolthaus et al., 2006). As a similarity measure, the CR takes values between 0 and 1, where a value of 0 indicates no functional dependence between the two images and a value of 1 can be interpreted as a high functional dependence. The correlation ratio is expressed as the variance of the conditional expectation of the target intensities given the reference intensities, relative to the variance of the target intensities (Roche et al., 1998a). As such, both the CR and the proposed measure based on LE maximize the variance of a set of projected intensities, the conditional expected intensities and the Laplacian embedded intensities, respectively. In the CR, the variance of the target intensities is split between the part that can be explained by the reference (which is maximized) and the part which cannot be explained (which is minimized). In the proposed LE-based measure, variance of the embedded intensities is split between the different eigenvalues. The variance which can be explained by the first eigenvalue is maximized, while the variance in the other eigenvalues is minimized. Note however that in the correlation ratio a functional relationship is assumed, meaning that for every input only one output value exists in the function. A manifold, employed in LE, is not bound by such restrictions and can be employed to represent more general intensity relationships, which is a fundamental point of difference. Furthermore, the correlation ratio is asymmetric with respect to the target and reference image, each taking a different role in the calculation of the measure (the target image being conditionalized on the reference). The proposed dissimilarity measure, however, is symmetric and each image in the group is considered equally.

6.2.4 Optimization

With the proposed dissimilarity measure being graph-based, its optimization can be challenging. Changing one edge in the graph induces a discrete, discontinuous (and thus non-differentiable) change in the measure value, making gradient-based optimization difficult. Alternatives in non-gradient-based optimization, such as an evolution strategy, would typically not guarantee optimality either and Klein et al. (2007) have previously shown that the convergence rate of such an evolution strategy is too low to be competitive. As such, in our experiments gradient-based optimization is performed. It should be noted that the following derivation for the gradient of the proposed measure should be considered as an approximation since it does not take into account possible changes in the topology of the graph.

The approach of de Leeuw (2007) was followed to calculate the derivative of an eigenvalue:

$$\frac{\partial \mathcal{D}_{LE}}{\partial \boldsymbol{\mu}} = \boldsymbol{v}_1^T \left(\frac{\partial \boldsymbol{\mathcal{L}}}{\partial \boldsymbol{\mu}} - \lambda_1 \frac{\partial \boldsymbol{D}}{\partial \boldsymbol{\mu}}\right) \boldsymbol{v}_1 \quad . \tag{6.6}$$

Herein is $\boldsymbol{v}_1$ the eigenvector associated to λ_1. To ease the notational burden we will drop the subscript on $\boldsymbol{v}_1$ and λ_1 hereafter and rewrite $\boldsymbol{W}$ from Eq. (6.2) as

$$\boldsymbol{W} = e^{-\frac{\widetilde{\boldsymbol{R}}}{t}} \quad . \tag{6.7}$$

Where $\widetilde{\boldsymbol{R}} = \tau(\boldsymbol{R})$ and τ the function is which sets all elements to zero if they are not connected with an edge in the graph. $\boldsymbol{R}$ is written as

$$\begin{aligned} \boldsymbol{R} &= \sum_{i=1}^{n} \boldsymbol{R}_i \\ &= \sum_{i=1}^{n} (\boldsymbol{m}_i \circ \boldsymbol{m}_i)\, \mathbf{1}^T + \mathbf{1}\, (\boldsymbol{m}_i \circ \boldsymbol{m}_i)^T - 2\boldsymbol{m}_i \boldsymbol{m}_i^T \end{aligned} \tag{6.8}$$

with $\mathbf{1}$ a column vector of size s with all elements equal to 1 and $\circ$ the Hadamard product. Its derivative can then be found to be:

$$\begin{aligned} \frac{\partial \widetilde{\boldsymbol{R}}}{\partial \mu} =& \tau \Bigg(2 \sum_{i=1}^{n} \left(\boldsymbol{m}_i \circ \frac{\partial \boldsymbol{m}_i}{\partial \mu} \right) \mathbf{1}^T + \mathbf{1} \left(\boldsymbol{m}_i \circ \frac{\partial \boldsymbol{m}_i}{\partial \mu} \right)^T - \\ & \boldsymbol{m}_i \left(\frac{\partial \boldsymbol{m}_i}{\partial \mu} \right)^T - \left(\frac{\partial \boldsymbol{m}_i}{\partial \mu} \right) \boldsymbol{m}_i^T \Bigg) \quad . \end{aligned} \tag{6.9}$$

Employing Eq. (6.7, 6.8) to rewrite $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{D}$ in Eq. (6.6) and expanding all products leads to

$$\begin{aligned}\frac{\partial \mathcal{D}_{LE}}{\partial \mu} =& \frac{-(1-\lambda)}{t} \sum_a v_a \sum_b \left(W_{ab} \frac{\partial \widetilde{R}_{ab}}{\partial \mu} \right) v_a - \\ & \frac{-1}{t} \sum_a v_a \sum_b \left(W_{ab} \frac{\partial \widetilde{R}_{ab}}{\partial \mu} \right) v_b \ .\end{aligned} \tag{6.10}$$

Introducing Eq. (6.9) herein results in

$$\begin{aligned}\frac{\partial \mathcal{D}_{LE}}{\partial \mu} =& \frac{-2}{t} \sum_{i=1}^{n} \sum_a \sum_b (1-\lambda) W_{ab} \\ & \Bigg[v_a^2 m_{i,a} + v_b^2 m_{i,a} - v_b^2 m_{i,b} - v_a^2 m_{i,b} \\ & - 2 v_a m_{i,a} v_b + 2 v_a v_b m_{i,a} \Bigg] \frac{\partial m_{i,a}}{\partial \mu} \ .\end{aligned} \tag{6.11}$$

In our implementation, the graph, needed to construct Eq. (6.2), was built using k nearest neighbors and a brute force approach which scales quadratically with respect to the number of coordinate samples, s. Note that other methods exist to construct the graph, most notably the approximate nearest neighbors algorithm (Arya et al., 1998) or a fixed-radius approach. To obtain the eigenvalues from Eq. (6.4), the Lanczos algorithm for sparse symmetric matrices was employed (Lanczos, 1950) which scales linearly with the number of coordinate samples and the average number of edges connecting to a node in the graph. Furthermore, the proposed algorithm requires the construction of several matrices (such as $\boldsymbol{W}$, $\boldsymbol{D}$, $\boldsymbol{\mathcal{L}}$ and $\widetilde{\boldsymbol{R}}$) of size $s \times s$, which would be troublesome in terms of memory consumption and computation time when the measure is calculated deterministically over the entire sampling domain, discretized to a voxel level. Registering a group of images of size $512 \times 512 \times 512$ voxels would require more than 32.000 TB for one such matrix at 16bit precision. It is clear that such an approach would be unfeasible and to resolve this issue, we propose to combine the LE-based measure with a stochastic optimization approach. Here, only a limited number of random coordinates samples, s, are considered in each iteration of the registration algorithm. As a result, the size of the matrices is reduced to a tractable level and significantly speeds up the calculation of the eigenvalues.

An adaptive stochastic gradient descent was employed where the number of random coordinate samples, s, was set to the default value of 2048, following Klein et al. (2009b). The number of iterations was fixed to the default value of 2000, after which convergence was assumed.

6.2.5 Hyperparameter settings

The application of Laplacian eigenmaps requires finetuning two hyperparameters, the number of nearest neighbors in the graph, k, and the free parameter in Eq. (6.2), t. Setting the nearest neighbors, k, to a smaller value leads to a more detailed characterization of the manifold, highlighting the outlier voxels which require the largest deformation. Care should be taken not to 'break' the graph into multiple subgraphs, as the Fiedler eigenvalue collapses to zero and is only defined for a connected graph. As such, the lowest number which does not break up the graph is a good starting point for finetuning the number of nearest neighbors, k. t can be rewritten to $2\sigma^2$ giving Eq. (6.2) its resemblance to a Gaussian kernel. As an alternative to finetuning the parameter, we propose to set this hyperparameter, σ^2, to 10 times the variance of $||\boldsymbol{m}^a - \boldsymbol{m}^b||^2$ in Eq. (6.2) and to recalculate it after a predefined number of iterations. This allows for nodes who are far apart to still influence the constructed manifold. Furthermore, dynamically setting the hyperparameter, σ^2, is necessary due to the changes in the manifold during the registration process. Note that alternative approaches exist to set this hyperparameter, t (Sharma et al., 2009).

Following the approach of Balci et al. (2007) we constrain the transformations such that the sum of all deformations for a given coordinate in the common sampling domain, Ω, is equal to zero

$$\frac{1}{n}\sum_{i=1}^{n} \mathcal{T}_{\boldsymbol{\mu}_i}(\boldsymbol{x}) = \boldsymbol{x}, \quad \forall \boldsymbol{x} \ . \tag{6.12}$$

6.3 Data and experiments

In this work, experiments were performed on four datasets. The first dataset featured synthetic data and the second dataset featured simulated data from the BrainWeb database (Cocosco et al., 1997; Collins et al., 1998). The third and fourth dataset included clinical data, MRI acquisitions of the brain in neonates (Wang et al., 2019) and MRI acquisitions of the knee (Balamoody et al., 2010; Williams et al., 2010).

6.3.1 Experimental settings

We compared the proposed LE-based measure with four other groupwise dissimilarity measures suited for multimodal registration problems. Accumulated pairwise estimates of mutual information (APEMI) compares every image in the group to every other image in terms of mutual information and sums these estimates (Wachinger and Navab, 2013). The average mutual information (AMI) compares every image in the group to the current average intensity image in terms of mutual information (Bhatia et al., 2007). The conditional template entropy (CTE) compares every image to the principal component image in terms of conditional entropy (Polfliet et al., 2018). Finally, the PCA2 measure was included, which is expressed as a weighted sum of the variance explained by the progressing eigenvalues of the intensity correlation matrix of the images (Huizinga et al., 2016). Aside from the choice of dissimilarity measure, all other experimental settings remained fixed for every experiment to avoid confounding factors. All experiments were performed single-threaded on the local high performance cluster.

A multiresolution approach was employed with deformable transformations based on cubic B-splines (Rueckert et al., 1999). The proposed dissimilarity measure was implemented as part of the software package `elastix` (Klein et al., 2010) and is made publicly available.

All deformable registrations were regularized with a groupwise bending energy (Polfliet et al., 2018). The hyperparameters related to the cost functions, such as the regularization weight, the number of histogram bins for entropy-based measures and the number of nearest neighbors in the graph, were fine-tuned with respect to the validation scores to obtain the best result for each (dis)similarity measure and dataset separately.

Hyperparameters unrelated to the cost function were set either to their default value, such as the number of random coordinate samples in the adaptive stochastic gradient descent optimizer, s, or empirically, such as the number of resolutions or the control point spacing. These hyperparameters were set to a constant value for all (dis)similarity measures in a given dataset. The employed values of the most important hyperparameters are presented in Table 6.1.

6.3.2 Synthetic dataset

First, an experiment performed by Roche et al. (1998b) was repeated for the groupwise measures under consideration. This experiment demonstrated the shortcoming of mutual information-based measures in images with bias fields,

Table 6.1: Summary of the registration hyperparameters employed in the different datasets.

Dataset	Number of images	Resolutions	Histogram bins	Grid spacing (mm)	Spatial samples	Iterations	Nearest neighbors
Synthetic	2	1	48	N.A.	1200/All	N.A.	400
BrainWeb	30-90	2	48	6.0	2048	2000	120
iSeg	10-46	7	32	4.0	2048	2000	65
MRI Knee	86	7	32	6.0	2048	2000	10

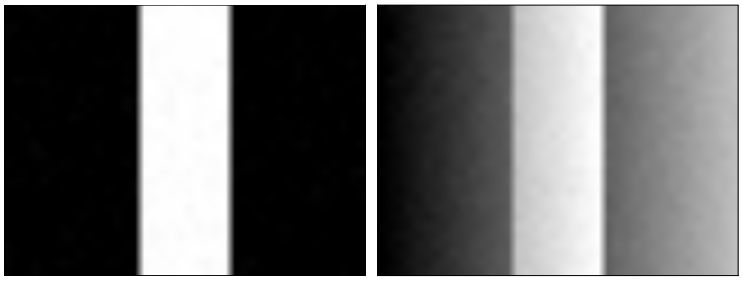


Figure 6.2: Images in the synthetic dataset with a bias field.

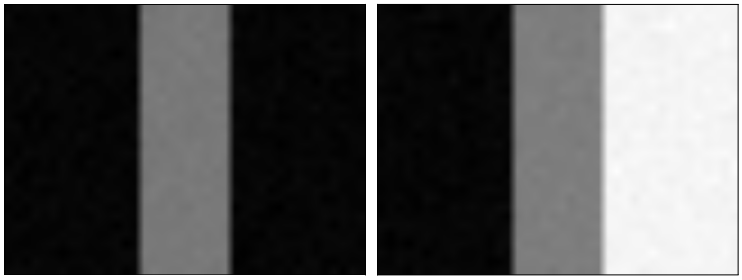


Figure 6.3: Images in the synthetic dataset with a non-linear intensity relationship.

which are typically present in MRI-based acquisitions. Two images of 30-by-40 pixels were progressively translated relative towards each other. The first image consisted of only two intensities (0 and 1), whereas the second image contained a continuous spectrum of intensities, resembling a bias field (Fig. 6.2). This hampers mutual information-based measures, as intensities are spread across different bins in the joint histogram, and alignment does not lead to a decrease in joint entropy.

Secondly, an experiment was performed where the first image consisted of only two intensities (0 and 0.5) and the second image contained three intensities (0, 0.5 and 1) to simulate a non-linear intensity relationship between the two images (Fig. 6.3). It is expected that measures based on linear dimensionality reduction will fail to model the non-linear intensity relationship and fail to indicate the correct optimum.

Thirdly, an experiment was performed to investigate the effect on computation time when increasing the number of images. The images from the second experiment were added to a group of increasing size (i.e. size 2, 5, 10, 20, 50, 100, 200, 500 and 1000) and the time to perform a single iteration was measured.

Due to the synthetic nature of the images (only three distinct intensities) and a limited amount of samples therein, all pixels of the images ($s = 1200$) were used to calculate the values of the (dis)similarity measures. Furthermore, σ^2 was set manually to be equal to the variance of $||\boldsymbol{m}^a - \boldsymbol{m}^b||^2$ instead of the automatic setting discussed in sec. 6.2.4. This proved to be necessary due to large discrete intensity differences in this synthetic example.

6.3.3 BrainWeb dataset

BrainWeb (Cocosco et al., 1997; Collins et al., 1998) allows for users to download simulated brain images of three MRI modalities (T1, T2 and PD). Examples of the images are shown in Fig. 6.4. For our experiments, additive Gaussian noise and a 40% intensity non-uniformity field were added to the images. The influence of an increasing number of images in the group was investigated for all (dis)similarity measures. T1 images were grouped and the number of images in the group was increased from 30 to 90 with a stepsize of 10. Separately, the influence of additional modalities in the group of images was investigated. 30 T2 and subsequently 30 PD images were added to a group of 30 T1 images, leading to a total of, respectively, 60 and 90 images.

Twenty random deformable transformations were synthesized for each separate group of images with a B-spline control point spacing of 8 voxels and the registrations attempted to recover the generated transformations with a control

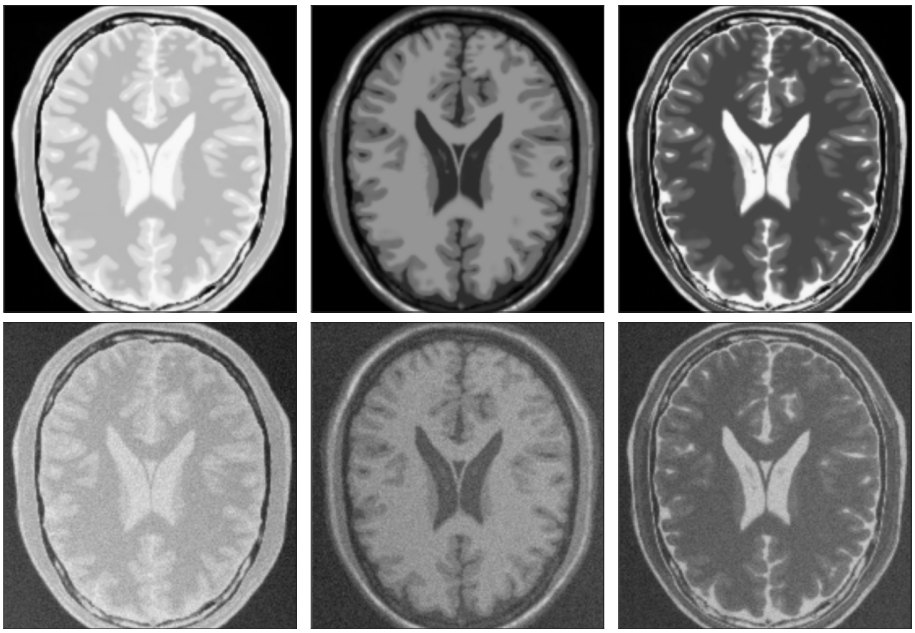


Figure 6.4: Top - Original BrainWeb images. Bottom - Noisy BrainWeb images. Left - PD modality. Center - T1 modality. Right - T2 modality.

point spacing of 6 voxels. The random transformations were confirmed to be invertible by verifying that the Jacobian determinant was positive everywhere. 40 landmarks were used to quantify the registration accuracy using the groupwise target registration error (gTRE):

$$\mathrm{gTRE}\left(\boldsymbol{\mu}\right) = \frac{1}{n}\sum_{i\neq r}^{n}\frac{1}{|P_i|}\sum_{j}^{|P_i|}||\mathcal{T}_{i,r}\left(\boldsymbol{p}_{i,j}\right) - \boldsymbol{p}_{r,j}|| \tag{6.13}$$

where r is the index of the reference image, P_i the collection of landmarks in the ith image, $\mathcal{T}_{i,r}$ the transformation that maps the coordinates from the ith image to the reference image and $\boldsymbol{p}_{i,j}$ the jth landmark from the ith image. The initial gTRE before registration is 1.51mm.

6.3.4 iSeg dataset

For the 2017 iSeg Challenge, 10 T1 and T2 pairs of neonatal brain images were made available with ground truth segmentations of the grey and white matter and cerebrospinal fluid, together with 13 T1 and T2 pairs without ground truth segmentations (Wang et al., 2019). The T1 images were acquired as sagittal slices with $TR/TE = 1900/4.38$ ms, a flip angle of 7° and a resolution of $1 \times 1 \times 1$ mm^3 and the T2 images as axial slices with a $TR/TE = 7380/119$ ms, a flip angle of 150° and a resolution of $1.25 \times 1.25 \times 1.25$ mm^3. These images were corrected for bias fields present in the MRI acquisitions. However, due to cortical organization and myelination of the white matter during the first two years of pediatric brain development, contrast inversion of gray and white matter can take place in MR imaging. This has led several authors to use multimodal (dis)similarity measures in registration tasks involved in brain atlas construction, even after intensity standardization algorithms have been applied (Bhatia et al., 2007; Blesa et al., 2016; Klein et al., 2009a; Xue et al., 2007).

Similar to the experiments performed with the BrainWeb dataset, we investigated the influence of increasing the number of images and modalities in the registrations. First, registrations were performed on 10 T1 images for which ground truth segmentations were available, then progressively including additional T1 images with a stepsize of 2. Secondly, starting from the same 10 T1 and 10 T2 images for which ground truth segmentations were available, additional T1-T2 pairs were included with a step size of 4 images (i.e. 2 pairs).

Validation of all registration results was performed on the 10 images where a ground truth segmentation was available, employing a groupwise Dice score (gDice) and a smoothness score (Smo):

$$\text{gDice}\left(\boldsymbol{\mu}\right) = \frac{1}{h}\frac{1}{n}\sum_{g=1}^{h}\sum_{i=1}^{n'}\frac{2|A_i^g \cap S_i^g \circ \mathcal{T}_{\boldsymbol{\mu}_i}|}{|A_i^g| + |S_i^g \circ \mathcal{T}_{\boldsymbol{\mu}_i}|} \tag{6.14}$$

$$\text{Smo}\left(\boldsymbol{\mu}\right) = \frac{1}{n}\sum_{i=1}^{n}\text{STD}\left(\left|\frac{\partial \mathcal{T}_{\boldsymbol{\mu}_i}}{\partial \boldsymbol{x}}\right|\right) \quad . \tag{6.15}$$

In Eq. 6.14, h is the number of structures which were annotated (in this case the grey and white matter and cerebrospinal fluid) and n' is the number of images for which segmentations are available. S_i^g is the segmentation of the gth structure in the ith image. $A_i^g = \lfloor \frac{1}{2} + \frac{1}{n'-1}\sum_{j\neq i}^{n'} S_j^g \circ \mathcal{T}_{\boldsymbol{\mu}_j} \rfloor$ is a majority

voted label atlas image of the gth annotated structure based on all transformed segmentations in the group except the segmentation that is being tested (to avoid any bias). The initial gDice score in this experiment before registration was 39%. In Eq. 6.15, 'STD' refers to the standard deviation taken over all coordinates, $\boldsymbol{x}$, in the image domain (Huizinga et al., 2016). The smoothness score can be employed to signal physically implausible or extreme deformations.

6.3.5 MRI knee dataset

86 T1 MRI images of the knee were made available by the OsteoArthritis Initiative with associated ground truth segmentations for the femoral cartilage, lateral and medial tibial cartilage and the lateral and medial meniscus (Balamoody et al., 2010; Williams et al., 2010). The images were acquired with sagittal slices and a double echo steady state sequence at a resolution of $0.36 \times 0.36 \times 0.7$ mm^3 and not preprocessed or corrected for intensity inhomogeneities commonly associated to MRI acquisitions. These intensity inhomogeneities can be considered to be similar to the problems faced in multimodal registration and are typically solved by using multimodal (dis)similarity measures (Klein et al., 2009a).

The experimental design is similar to the experiments in the iSeg and BrainWeb dataset. The influence of an increasing number of images was investigated by performing registrations with 36, 46, 56, 66, 76 and 86 images. Similar to the iSeg dataset, validating the obtained transformations was achieved with the gDice (Eq. 6.14) of the five available anatomical structures and the smoothness score (Eq. 6.15). The intial gDice score before registration was 0% due to poor initialization. To be able to directly compare the validation results across registrations with different number of images, the same 36 ground truth segmentations were used across all validations, even when more segmentations were available.

Additionally, to investigate the impact the hyperparameters can have on the registration accuracy and to verify the effectiveness of the automatic hyperparameter estimation of t, a hyperparameter robustness experiment was performed. Here, the values of the two hyperparameters of the proposed dissimilarity measure, t and k, were varied. These registrations were performed without regularization to avoid possible confounding effects.

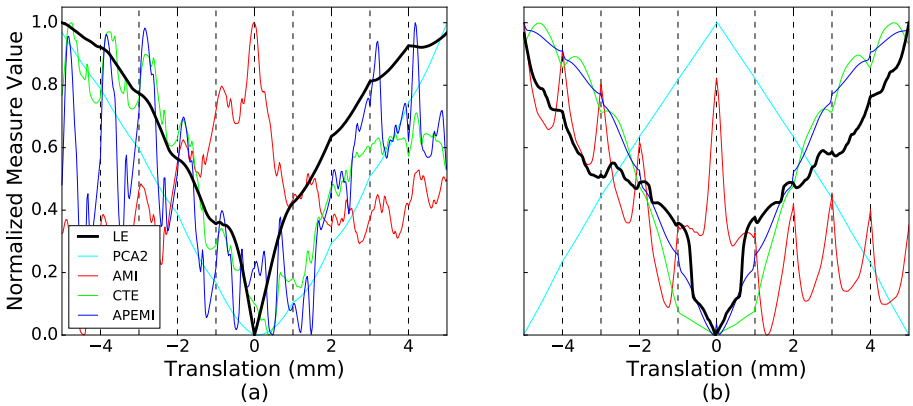


Figure 6.5: The behavior of the different (dis)similarity measures with images from the Synthetic dataset (a) including a bias field and (b) including a non-linear intensity relationship. The measure values are normalized between 0 and 1 for clarity. The dashed vertical lines indicate integer voxel displacements.

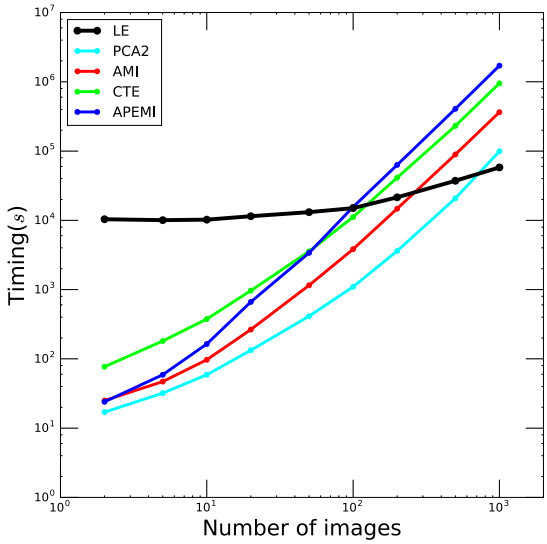


Figure 6.6: The computation time of a single iteration for the investigated dissimilarity measures. Note the log-log scale.

6.4 Results

6.4.1 Synthetic

The behavior of the different (dis)similarity measures for the Synthetic dataset is given in Fig. 6.5 with the measure values normalized between 0 and 1 for clarity. All measures under investigation are implemented as dissimilarity measures, i.e. minimizing the negative mutual information instead of maximizing it. As such, the global optimum is achieved when the normalized measure value reaches zero.

The proposed measure based on LE was able to identify the correct minimum in both cases. In the experiment involving images with a bias field, all entropy-based measures (AMI, CTE and APEMI) failed to recognize the correct minimum. In the experiment involving images with a non-linear intensity relationship, the PCA-based measure (and AMI) failed to indicate the correct minimum. It should be noted that all measures, except PCA2, suffered from intensity interpolation artifacts at non-integer voxel displacements similar to the experiment performed in Roche et al. (1998b).

Furthermore, in the third experiment, where we explored the time a single iteration requires, we can see in Fig. 6.6 that the LE-based measure has a completely different scaling behaviour when increasing the number of images. Increasing the size of the group of images 500-fold (from 2 to 1000) increased the computational time approximately 6-fold ($10 \times 10^3 s$ to $58 \times 10^3 s$), whereas APEMI increased the computation time 72000-fold.

6.4.2 BrainWeb

Results on the accuracy and computation time for the BrainWeb dataset are given in Fig. 6.7(a) and 6.8(a).

In the monomodal registrations involving only T1 images of the BrainWeb phantom, PCA2 performed best in terms of accuracy with a gTRE of 1.10 mm to 0.90 mm for registrations of 30 and 90 images, respectively; where LE scored second best with a gTRE of 1.23 mm to 1.25 mm. For the registrations including multimodal data, PCA2 performed best with a final gTRE of 0.94 mm and LE performed second best with a gTRE of 1.40 mm. In general, entropy-based measures performed poorly in terms of accuracy compared to measures based on dimensionality reduction and even deteriorated the initial alignment with an intial gTRE of 1.51 mm.

In terms of computational cost, registrations with LE showed the best scaling behaviour with the computation time only increasing twofold when the number

of images in the group was increased from 30 to 90. Other measures showed close to a tenfold increase. However, only when 90 images were used in the registration, the computation time of LE was equivalent or better than the other (dis)similarity measures.

6.4.3 iSeg

In Fig. 6.7(b), 6.8(b) and 6.9(a), the results on the accuracy, computation time and smoothness for the iSeg dataset are provided. In Fig. 6.10 vizualizations of the deformations fields are shown.

Both APEMI and LE demonstrated relatively stable registration accuracy for an increasing number of images with similar registration accuracy. Their gDice scores remained stable from 78.6% and 77.4% for 10 images in the group to 77.9% and 77.8% for 46 images in the group, for APEMI and LE respectively. The registration accuracy of template-based methods (AMI and CTE) suffered from the inclusion of multimodal images in the group. Despite the multimodal nature of the T1 and T2 images, PCA2 performed adequately compared to the other (dis)similarity measures.

The transformations for PCA2 were most smooth compared to other approaches, whilst APEMI and LE performed comparable to each other and template-based methods performed worst.

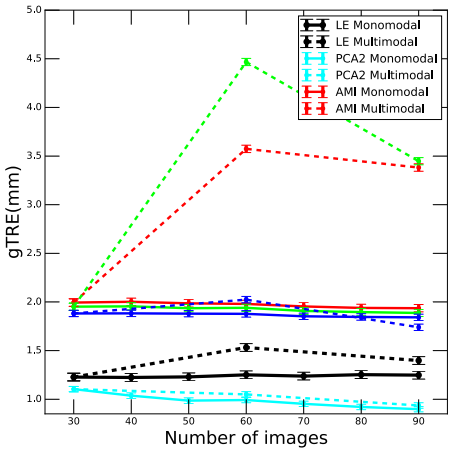
In terms of computation time, registrations with the reference measures (AMI, CTE, APEMI and PCA2) performed comparable, showing large increases in their computation time as the number of images in the group increased. Whereas registrations with LE are associated with a high cost for small groups of images, they showed little dependence on the computation time of the number of images in the group, becoming competitive in terms of efficiency from 45 images upwards.

6.4.4 MRI knee

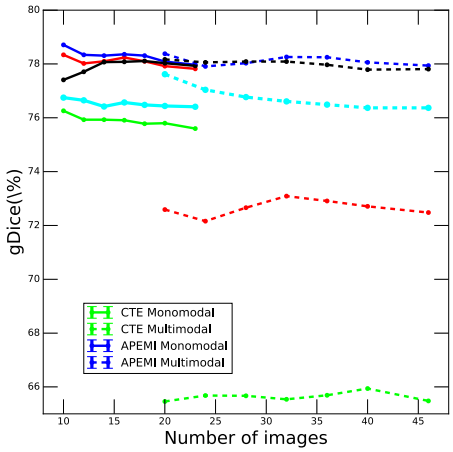
The gDice scores, computation times and smoothness scores for the experiment with MRI Knee data are given in Fig. 6.7(c), 6.8(c) and 6.9(b). In Fig. 6.11 the group of images is displayed as a concatenation before and after registration.

The results for the gDice show comparable performance for all (dis)similarity measures for all image group sizes (within 1% gDice).

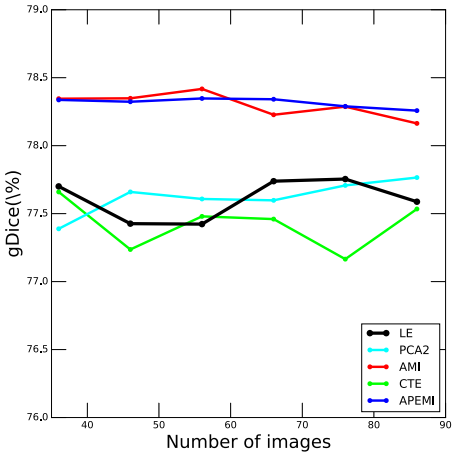
In terms of the smoothness score, the proposed dissimilarity measure performed best. For all methods, a decreasing trend could be noticed when more images were included in the registration.



(a) BrainWeb

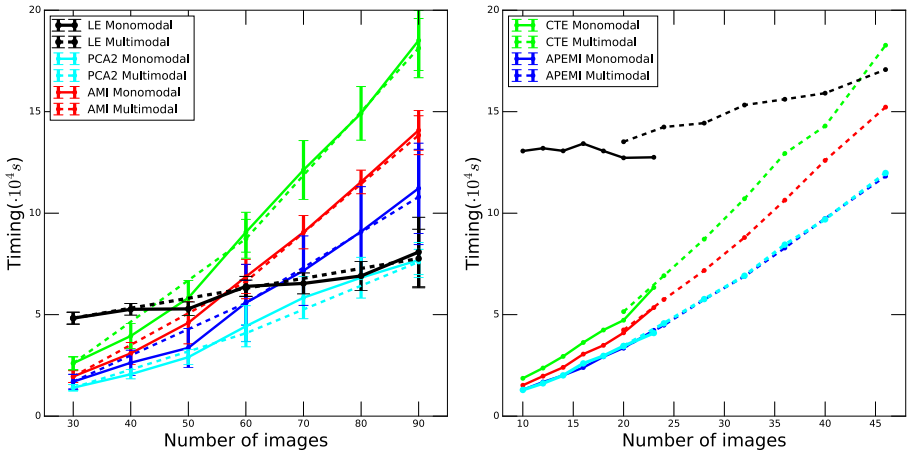


(b) iSeg



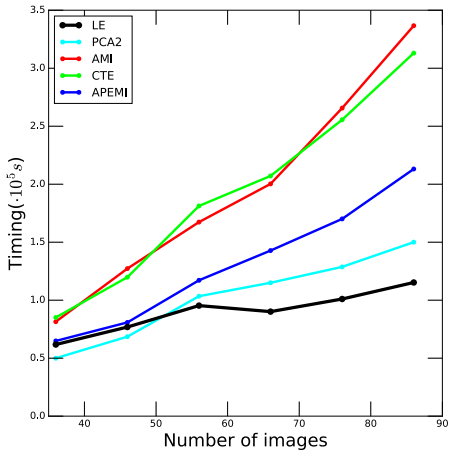
(c) MRI Knee

Figure 6.7: Registration accuracy results as a function of the number of images in the registration for the (a) BrainWeb, (b) iSeg and (c) MRI Knee dataset. (a) gTRE results (lower is better) for the BrainWeb dataset. The errorbars indicate the standard deviation over the 20 repeated registrations. 'Monomodal' refers to registrations including T1 images only and 'Multimodal' refers to registrations which also includes T2 and/or PD images. (b) gDice scores (higher is better) for the iSeg dataset. 'Monomodal' refers to registrations including T1 images only and 'Multimodal' refers to registrations which also included T2 images. (c) gDice scores (higher is better) for the MRI Knee dataset.



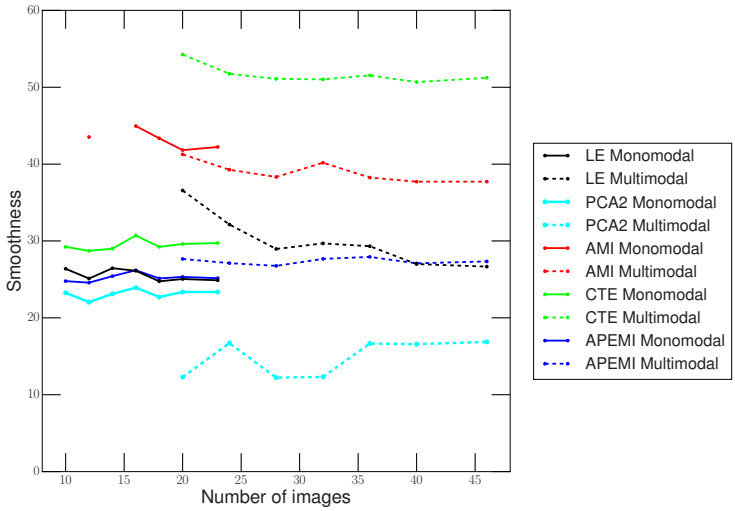
(a) BrainWeb

(b) iSeg

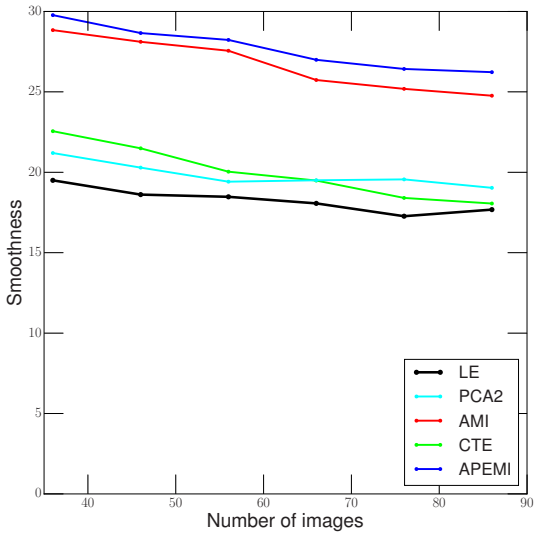


(c) MRI Knee

Figure 6.8: Computation times (lower is better) as a function of the number of images in the registration for the (a) BrainWeb, (b) iSeg and (c) MRI Knee dataset. In (a) the errorbars indicate the standard deviation over the 20 repeated registrations, 'Monomodal' refers to registrations including T1 images only and 'Multimodal' refers to registrations that also included T2 and/or PD images. In (b) 'Monomodal' refers to registrations including T1 images only and 'Multimodal' refers to registrations that also included T2 images.

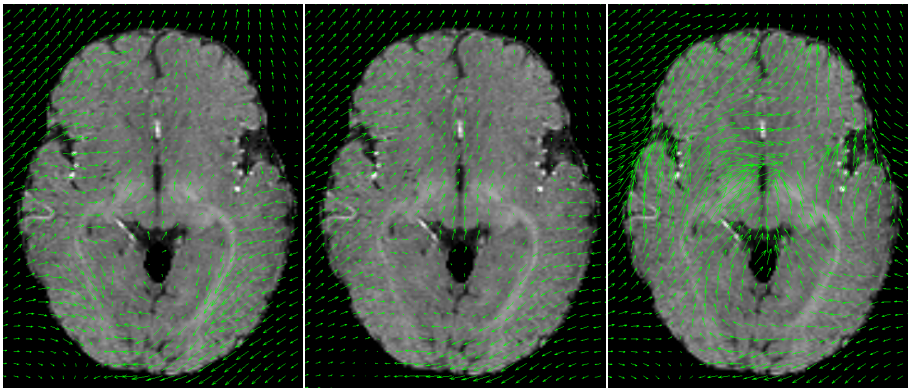


(a) iSeg

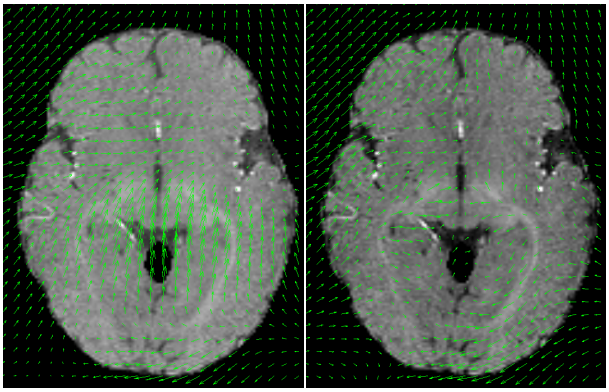


(b) MRI Knee

Figure 6.9: Registration smoothness results (lower is better) as a function of the number of images in the registration for the (a) iSeg and (b) MRI Knee dataset. In (a) 'Monomodal' refers to registrations including T1 images only and 'Multimodal' refers to registrations that also included T2 images. Note that for AMI some points could not be plotted on the graph. This was due to extreme bending at the edge of the ROI which created extreme outlier values.



(a) LE (b) PCA2 (c) AMI



(d) CTE (e) APEMI

Figure 6.10: Axial slice of the iSeg dataset overlayed with the resulting deformation field for the different measures under study.

Large differences can be observed for the computation times. Herein, LE scored best and required 32 hours to complete the registration when all 86 images are employed in the registration. As a result a speedup of factor 2-3 is achieved compared to the other measures. In addition to the best absolute performance in terms of computation times, the proposed measure also showed the best scaling behaviour.

For the hyperparameter robustness experiment the results are shown in Fig. 6.12. Limited impact on the registration accuracy is noted over the tested range of hyperparameter values. The automatic hyperparameter estimation for t was effective to both simplify the hyperparameter optimization and increase the accuracy of the method.

6.5 Discussion

In the experiments based on synthetic data, the correct optimum was found by the proposed method based on Laplacian Eigenmaps, showing its potential to be applied to various registration problems. The experiments on synthetic data highlight that LE can be utilized both for multimodal data with non-linear intensity relationships, where PCA-based measures would fail; and for data with complex intensity profiles caused by bias fields, where entropy-based measures would fail. Additionally, in the third experiment we highlight the favorable scaling behaviour of the computation time with respect to the number of images in the registration for the proposed measure based on LE.

The experiments based on the BrainWeb, iSeg and MRI Knee dataset showed consistent results. First, in terms of registration accuracy, the LE measure performed second best (after PCA2) in the experiments on the BrainWeb dataset, achieves the best results (together with APEMI) in the experiments based on the iSeg dataset and performs similar to the other dissimilarity measures in the experiment on the MRI Knee dataset. Secondly, in all three datasets the proposed methodology showed an increase in computation times which was considerably less with an increasing number of images compared to other (dis)similarity measures. These results imply an efficient computational approach for large groups of images as we previously highlighted in section 6.2.4. Note, however, that some increase in the computation times with an increasing number of images is to be expected due to other components of the registration algorithm with such a dependency (e.g reading the images to memory or the calculation of the regularizer).

For the smoothness results we can observe a trade-off between a more ac-

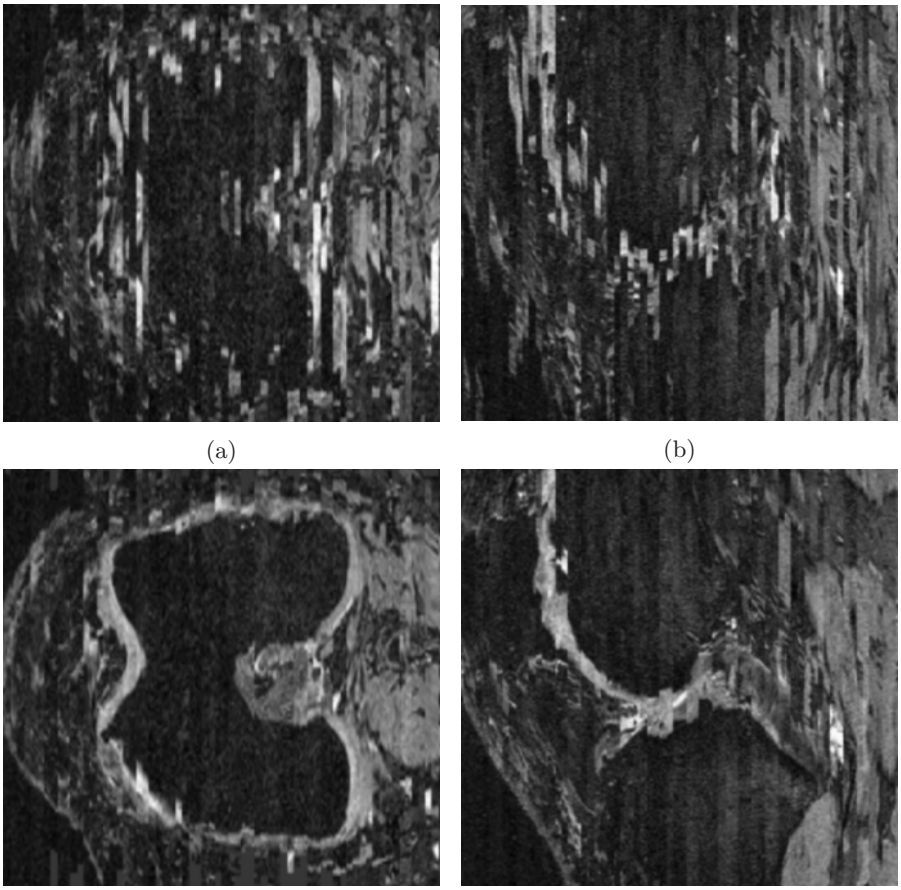


Figure 6.11: Partial concatenation of all images used in the MRI knee dataset. Each image in the group is represented in the concatenation image in four to five sequential slices along the coronal axis to highlight the remaining intersubject differences in the anatomical structures (a) Axial and (b) sagittal view before registration. (c) Axial and (d) sagittal view after registration.

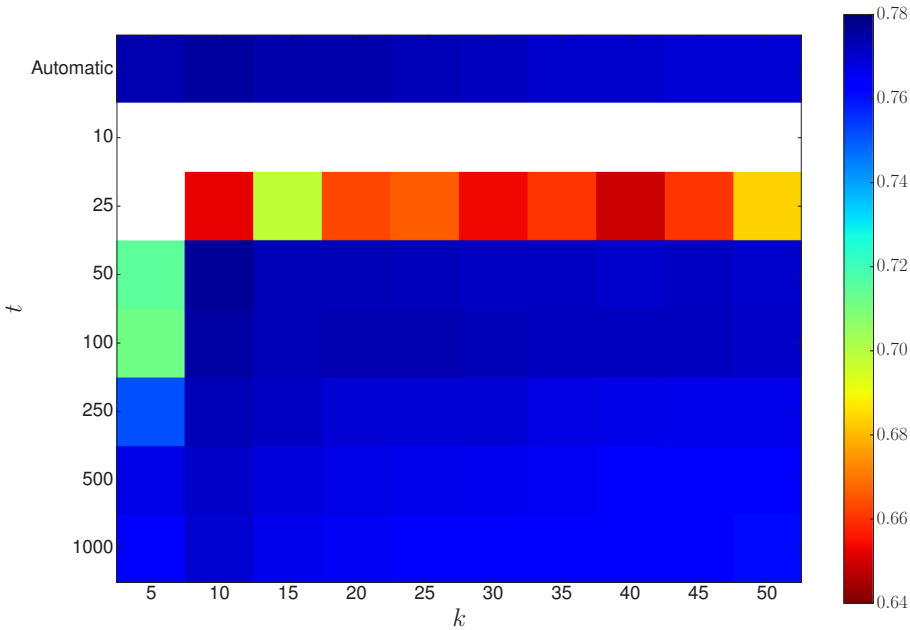


Figure 6.12: Hyperparameter robustness results for registrations on the MRI knee dataset with varying hyperparameters, t and k. The colour coding represent the gDice scores and the white squares indicate failed registrations. The top row indicates the automatic hyperparameter estimation proposed in this work.

curate, but simultaneously a more complex (i.e. non-smooth) transformation in Fig. 6.9(b). This trade-off holds in Fig. 6.9(a) if the template-based methods (AMI and CTE) are excluded. This is qualitatively confirmed in Fig. 6.10. The most complex deformation fields were generated from AMI and CTE, while PCA2 produced the simplest deformation field. It should be noted that producing smooth transformations might not necessarily be a desirable property of a (dis)similarity measure. For example, in cases where sliding motion occurs (such as in the thoracic or abdominal region) or in inter-subject registration where certain structures might not co-occur in both patients (such as implants or lesions). This was not the case in this work.

It is important to emphasize that in these datasets, neither the computa-

tion times, the transformation smoothness nor the registration accuracy suffered from including additional images in the registration for the proposed methodology. This property could aid in the unbiased construction of population atlases where only a limited number of annotations are available. Including additional and possibly unannotated images could lead to a closer representation of the true population average to which the annotated images could be mapped, without the burden of extra computation time or reducing the smoothness or accuracy of the registration result. This was illustrated in Fig. 6.13. Here, the label and intensity atlas images were constructed from the 10 image sets for which the annotations were available in the iSeg dataset. In the top row, the source intensity and label images were transformed to the atlas space with the result from the groupwise registration where the intensity images were utilized for which annotations were available ($n = 10$). In the bottom row, all intensity images ($n = 23$) were incorporated in the groupwise registration and the transformations to the atlas space for the 10 labeled image sets were extracted and applied before constructing the atlas images. It is clear that a different resulting atlas was obtained, potentially more representative of the population mean, due to the use of more images in the analysis. Herein, a registration technique for which the registration accuracy (or timing) is not hampered by the number of images in the group is especially important.

PCA2 has performed on par with other multimodal (dis)similarity measures in all experiments with non-synthetic data despite the multimodal nature of some of the included images. In the BrainWeb dataset, we registered T1, T2 and PD images, whereas T1 and T2 images were registered in the iSeg dataset. It is plausible to assume that given the similarity in the imaging modalities, an almost linear subspace exists that PCA-based measures can exploit to drive the registration. The linear assumption of PCA-based measures obtains a simpler embedding or representation compared to the non-linear embedding achieved with Laplacian Eigenmaps. With more degrees of freedom to obtain the embedding, Laplacian Eigenmaps may overfit and embed noise, especially when a simpler embedding exists. This could explain the superior performance of PCA-based measures compared to the proposed LE-based measure in the BrainWeb dataset. Additionally, it should be noted that PCA2 uses all eigenvalues of the intensity correlation matrix to compute the dissimilarity measure. Exploiting other eigenvalues in the LE-based measure, besides the well-known Fiedler eigenvalue, could improve the registration accuracy and would be an interesting direction for future research.

In this work we compared different dissimilarity measures in terms of registration accuracy with a gTRE or Dice score. However, other measure prop-

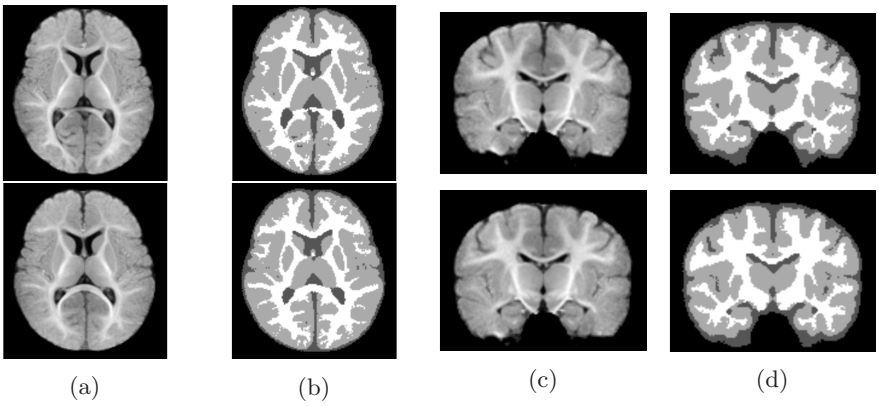


Figure 6.13: (a) Axial and (c) coronal view of the intensity atlas images constructed using intensity averaging. (b) Axial and (d) coronal view of the corresponding label atlas images constructed using majority voting. In the top row, the source intensity and label images were transformed to the atlas space with the result from the groupwise registration where the intensity images were utilized for which annotations were available ($n = 10$). In the bottom row, all intensity images ($n = 23$) were incorporated in the groupwise registration and the transformations to the atlas space for the 10 labeled image sets were extracted and applied before constructing the atlas images.

erties such as the capture range, distinctiveness of the optimum or risk of non-convergence (Skerl et al., 2006) could be of interest and insightful as well, and should be studied in future research.

It is interesting to discuss the proposed measure based on Laplacian eigenmaps and its similarities to entropic graphs for groupwise registration. Both approaches are graph-based, where the graph is constructed in the joint intensity space. Where LE tries to minimize the inverse of the explained variance with the first eigenvector of the Laplacian, entropic graphs try to minimize some length descriptor (minimal spanning tree length, k-nearest neighbors graph length, ...) of the graph. With the Fiedler eigenvalue closely related to the average distance of the nodes in the graph (Mohar et al., 1991), it is expected both methods are closely related. The relationship between both methods remains to be fully investigated.

Given the wide range of registration problems in which the correlation ratio has been used, and the similarities between the proposed measure and the

correlation ratio, it would be interesting to expand the validation to additional modalities such as ultrasound and CT in the future.

6.6 Conclusion

In this work we have proposed a novel dissimilarity measure for multimodal groupwise registration. The method is based on Laplacian Eigenmaps, a non-linear dimensionality reduction technique, and is employed to capture complex non-linear intensity relationships. The dissimilarity is expressed as the magnitude of the Fiedler eigenvalue, the second generalized eigenvalue of the Laplacian of the graph constructed in the joint intensity space. The measure, combined with the automatic hyperparameter estimation and stochastic gradient descent, showed improved or equivalent performance compared to other state-of-the-art groupwise (dis)similarity measures in terms of registration accuracy, in addition to an efficient scaling of the computational complexity with respect to the number of images in the group.

CHAPTER 7

Discussion and conclusion

7.1 Main findings

This thesis contains two main contributions. We compare groupwise registration to repeated pairwise registration in several experiments with respect to registration accuracy and transitivity, a measure which can be interpreted as the consistency of the transformations in a groupwise setting. Secondly, we fill a gap in current literature on efficient (dis)similarity measures for multimodal groupwise registration.

In **Chapter 3**, registrations were performed in the head and neck region of patients with oral squamous cell carcinoma. For this clinical application, we concluded that a conventional pairwise, rigid registration method is the recommended approach. Furthermore, we showed improved performance in terms of registration accuracy for the symmetric non-rigid transformation approach compared to the asymmetric approach, although the differences did not reach statistical significance. In other works the impact on registration accuracy was found to be significant when applying a symmetric registration approach Aganj et al. (2017); Lorenzi et al. (2013). Combining our findings with those in literature, we conclude that the choice of transformation strategy (symmetric vs. asymmetric) can have an impact on the registration accuracy. This is of importance to groupwise registration approach as they are often combined with symmetric transformations to reduce the methodological bias in the registration approach.

In **Chapter 4**, an investigation was performed on different template images in groupwise registration based on mutual information. We showed that the entropy of the template image can have a counter-productive contribution to the global measure value. Additionally, in this chapter we performed an initial

comparison in terms of registration accuracy between repeated pairwise and groupwise registration approaches. We showed that equivalent performance in terms of registration accuracy can be achieved with both approaches.

In **Chapter 5**, a novel similarity measure was introduced for multimodal groupwise registrations. The conditional template entropy measures the average of the pairwise similarity of each image of the group and a template image, which is constructed with the use of principal component analysis. The pairwise similarity is measured with the conditional entropy. The computational complexity of the novel measure scales linearly with respect to the number of images in the group. We furthermore showed improved or equivalent performance in terms of accuracy compared to other state-of-the-art (dis)similarity measures for multimodal groupwise registration and compared to multimodal repeated pairwise registrations. We showed that groupwise registrations vastly outperformed repeated pairwise registrations in terms of transitivity, which measures the transitive property of the transformations in a group of images and can be interpreted as a measure for the consistency of the transformations in a groupwise setting. These results are consistent with present literature (Geng et al., 2009; Metz et al., 2011) and this property can be identified as one of the advantages of groupwise approaches compared to repeated pairwise approaches.

In **Chapter 6**, to further improve on the efficiency of multimodal groupwise registration for large groups of images, we proposed a novel dissimilarity measure which is especially adept at registering such large groups of images. Laplacian eigenmaps were employed, a technique that achieves non-linear dimensionality reduction. The measure was formulated as the second smallest eigenvalue of the generalized eigenvalue problem posed in the description of Laplacian eigenmaps. The measure scales quadratically with the number of image samples employed in the measure calculation. It is clear that such an approach would be unfeasible when considering all image samples and to resolve this issue, we proposed to combine the LE-based measure with a stochastic optimization approach, reducing the calculations to a tractable level. We showed a favourable scaling of the measure in terms of computational time with respect to the number of images in the group and showed equivalent or improved performance in terms of registration accuracy compared to state-of-the-art multimodal groupwise (dis)similarity measures.

7.2 Future perspectives

In this work all registrations were performed in `elastix`, a modular and open-source library developed for medical image registration (Klein et al., 2010). Initially designed for pairwise image registrations, `elastix` can also be employed for groupwise registrations (Metz et al., 2011) despite a number of practical, software implementation related drawbacks. All images in the group that need to be registered are loaded into memory twice and must have the same origin, voxel spacing and size, requiring them to be resampled prior to the registration in some cases. Furthermore, the transformations in groupwise registrations in `elastix` are formulated from the domain of the common reference space to the domain of the reference spaces of the individual images in the group. Such a formulation might be suboptimal (Allassonnière et al., 2007; Ma et al., 2008). Efficient, modular and open-source software for groupwise registration has yet to emerge and would increase its adoption rate in the image registration community and clinical routine. Currently, the registration of large groups of images is limited to powerful machines capable to fit the entire group into its memory.

Additionally, such a modular and open-source tool would allow for researchers to evaluate newly proposed methods to existing state-of-the-art methods more easily. Rapid advances in AI research have shown that the focus on open source and open data has a beneficial effect on research and innovation output. Open (e-)health data has historically, and more recently with GDPR, been difficult to achieve (Culnane et al., 2017). More recently however, advances have been made to ensure privacy through novel methodologies (secure multi-party computation, homomorphic encryption, differential privacy, ...), proper infrastructure and education of the data owners and data processors. Which leads to more and more (open) data being available to transparently and reproducibly validate and benchmark novel methods against in the medical domain.

The increasing availability of health data drives some interesting research questions. Does the inclusion of more images in a groupwise registration result in an improved registration result as we illustrated in Fig. 2.3? In other words, does the additional data and the information it might carry (possibly resulting in a better defined global optimum) outweigh the additional algorithmic complexity? We could not find a clear answer to this question in present literature. In Chapter 6, we have illustrated that for some methods the registration accuracy and transformation smoothness might be impacted, both favourably and unfavourably. However, a thorough study is required to answer this research question. Note that given the computational scaling of some registration algo-

rithms, increasing the number of images can have a detrimental effect on the computational time required to perform the registration. In such cases a trade-off between accuracy and computational timing may be required. Furthermore, it is interesting to discuss this research question in parallel to the peaking paradox or phenomenon (occasionally seen in pattern recognition research). This paradox states that in a pattern recognition classifier (or regressor), the recognition accuracy initially increases when more features are added. However, once enough features have been added the accuracy peaks and deteriorates thereafter.

Several works show that groupwise registrations (Metz et al., 2011; Huizinga et al., 2016) obtain smoother transformations compared to repeated pairwise approaches. In this work we did not investigate what the driving force is behind this regularizing effect. However, it would be interesting from both a theoretical and practical perspective to investigate the relationship of this smoothing of the transformation further, especially its relationship to other explicit regularization terms such as the bending energy penalty or a localized rigidity penalty. Does explicit regularization influence the transformation smoothness in the same or in a different manner? Is an explicit regularization term required in groupwise registration? Furthermore, does this smoothing effect increase or grow when more images are included in the groupwise registration? And should groupwise approaches be avoided when smooth transformations are undesirable?

In recent years, major advances in learning-based approaches for image registration have shown promising results, especially those based on convolutional neural networks. In such approaches registering images can be performed in seconds instead of minutes or even hours, although at the cost that extensive prior training is required. Such learning-based approaches can be subdivided in supervised and unsupervised approaches. Where the former approach relies on ground truth transformations to be available at the time of learning (which are notoriously difficult to attain) (Cao et al., 2017; Sokooti et al., 2017), the latter typically relies on a synthetically constructed loss function based on segmentations or (dis)similarity measures such as the mean squared differences, the cross-correlation or another differentiable measure (Balakrishnan et al., 2019; de Vos et al., 2019; Sedghi et al., 2018; Yu et al., 2020; Lei et al., 2020). A model is trained to generate a smooth displacement field at every voxel by minimizing said loss function between the target image and the transformed source image after the application of the displacement field. For those unsupervised learning-based approaches based on (dis)similarity measures, the proposed measures from Chapters 5 and 6 can be plugged in and serve as a loss function. This could potentially speed-up registration even further while producing results, with similar or even improved accuracy. As such, developments in both

learning and non-learning based approaches could parallelize efforts to improve registration results.

Specifically in the context of groupwise image registration, it is interesting to discuss the differences and similarities that exist between non-learning and learning-based approaches. Consider the case where a large group of images needs to be registered. A learning-based approach performs prior training on all images in the large group, extracting valuable feature representations from the group, and compresses it into an efficient model to perform image registration. A non-learning-based groupwise approach, on the other hand, would include the entire group of images in the optimization problem without any compression which would require significantly greater computational resources. However, considering all images simultaneously at runtime has the advantage that was pointed out in Fig. 2.3. The ability to register two images, without any common features between them, through confounding features in a third image remains a compelling methodological advantage uniquely attributed to groupwise image registration. As such, a combined approach where learning-based groupwise registration is applied, would carry both the advantage of efficiency and possible confounding image features.

More generally, outside the medical research field, learning-based computer vision for pattern recognition could potentially benefit from efficient multimodal groupwise registrations as well. When faced with too much variance in the data, the images could be aligned to reduce the input data variability, for example in the case of facial recognition such alignment is common practice.

7.3 General conclusion

This thesis deals with advances in groupwise image registration, which is of growing interest to the community given the increasing availability of medical imaging data, both at the individual and the population level. We have shown that groupwise registrations achieve comparable results in terms of registration accuracy compared to repeated pairwise approaches and that groupwise approaches are superior in terms of the transitive error due to simultaneous estimation of all transformations. We identified a lack of efficient (dis)similarity measures for multimodal groupwise registration in existing literature and proposed two novel measures as a solution: the conditional template entropy, an entropy-based measure with a template image constructed with principal component analysis, and a measure based on Laplacian eigenmaps. For these measures specifically, we achieved equivalent or superior results in terms of registration

accuracy compared to other state-of-the-art registration methods.

Bibliography

Adaszewski, S., Dukart, J., Kherif, F., Frackowiak, R., Draganski, B., Initiative, A.D.N., et al., 2013. How early can we predict Alzheimer's disease using computational anatomy? Neurobiology of aging 34, 2815–2826.

Aganj, I., Iglesias, J.E., Reuter, M., Sabuncu, M.R., Fischl, B., 2017. Mid-space-independent deformable image registration. NeuroImage 152, 158–170.

Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, J.V., Rueckert, D., 2009. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. Neuroimage 46, 726–738.

Allassonnière, S., Amit, Y., Trouvé, A., 2007. Towards a coherent statistical framework for dense deformable template estimation. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69, 3–29.

Arya, S., Mount, D.M., Netanyahu, N.S., Silverman, R., Wu, A.Y., 1998. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. Journal of the ACM (JACM) 45, 891–923.

Avants, B.B., Tustison, N.J., Stauffer, M., Song, G., Wu, B., Gee, J.C., 2014. The Insight ToolKit image registration framework. Frontiers in neuroinformatics 8, 44.

Azampour, M.F., Ghaffari, A., Hamidinekoo, A., Fatemizadeh, E., 2014. Manifold learning based registration algorithms applied to multimodal images,

in: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE. pp. 1030–1034.

van Baar, G.J., Forouzanfar, T., Liberton, N.P., Winters, H.A., Leusink, F.K., 2018. Accuracy of computer-assisted surgery in mandibular reconstruction: A systematic review. Oral Oncology 84, 52–60.

Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V., 2019. Voxelmorph: a learning framework for deformable medical image registration. IEEE transactions on medical imaging .

Balamoody, S., Williams, T.G., Waterton, J.C., Bowes, M., Hodgson, R., Taylor, C.J., Hutchinson, C.E., 2010. Comparison of 3T MR scanners in regional cartilage-thickness analysis in osteoarthritis: a cross-sectional multicenter, multivendor study. Arthritis research & therapy 12, R202.

Balci, S.K., Golland, P., Shenton, M., Wells, W.M., 2007. Free-form B-spline deformation model for groupwise registration, in: Medical image computing and computer-assisted intervention: MICCAI... International Conference on Medical Image Computing and Computer-Assisted Intervention, NIH Public Access. p. 23.

Barroso, E.M., Ten Hove, I., Schut, T.C.B., Mast, H., van Lanschot, C.G., Smits, R.W., Caspers, P.J., Verdijk, R., Hegt, V.N., de Jong, R.J.B., et al., 2018. Raman spectroscopy for assessment of bone resection margins in mandibulectomy for oral cavity squamous cell carcinoma. European Journal of Cancer 92, 77–87.

Baumgartner, C.F., Kolbitsch, C., Balfour, D.R., Marsden, P.K., McClelland, J.R., Rueckert, D., King, A.P., 2014. High-resolution dynamic MR imaging of the thorax for respiratory motion correction of PET using groupwise manifold alignment. Medical image analysis 18, 939–952. doi:10.1016/j.media.2014.05.010.

Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., 2008. Speeded-up robust features (surf). Computer vision and image understanding 110, 346–359.

Belkin, M., Niyogi, P., 2003. Laplacian eigenmaps for dimensionality reduction and data representation. Neural computation 15, 1373–1396. doi:10.1162/089976603321780317.

Bhatia, K.K., Hajnal, J., Hammers, A., Rueckert, D., 2007. Similarity metrics for groupwise non-rigid registration, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007. Springer, pp. 544–552. doi:10.1007/978-3-540-75759-7_66.

Bhatia, K.K., Hajnal, J.V., Puri, B.K., Edwards, A.D., Rueckert, D., 2004. Consistent groupwise non-rigid registration for atlas construction, in: Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on, IEEE. pp. 908–911. doi:10.1109/isbi.2004.1398686.

Binahmed, A., Nason, R.W., Abdoh, A.A., 2007. The clinical significance of the positive surgical margin in oral cancer. Oral oncology 43, 780–784.

Blatt, S., Ziebart, T., Krüger, M., Pabst, A.M., 2016. Diagnosing oral squamous cell carcinoma: How much imaging do we really need? A review of the current literature. Journal of Cranio-Maxillofacial Surgery 44, 538–549.

Blesa, M., Serag, A., Wilkinson, A.G., Anblagan, D., Telford, E.J., Pataky, R., Sparrow, S.A., Macnaught, G., Semple, S.I., Bastin, M.E., et al., 2016. Parcellation of the healthy neonatal brain into 107 regions using atlas propagation through intermediate time points in childhood. Frontiers in neuroscience 10.

du Bois dAische, A., De Craene, M., Geets, X., Grégoire, V., Macq, B., Warfield, S.K., 2007. Estimation of the deformations induced by articulated bodies: Registration of the spinal column. Biomedical Signal Processing and Control 2, 16–24.

Bookstein, F.L., 1989. Principal warps: Thin-plate splines and the decomposition of deformations. IEEE Transactions on pattern analysis and machine intelligence 11, 567–585.

Brock, K.K., Dawson, L.A., Sharpe, M.B., Moseley, D.J., Jaffray, D.A., 2006. Feasibility of a novel deformable image registration technique to facilitate classification, targeting, and monitoring of tumor and normal tissue. International Journal of Radiation Oncology* Biology* Physics 64, 1245–1254.

Brown, B.J., Toler-Franklin, C., Nehab, D., Burns, M., Dobkin, D., Vlachopoulos, A., Doumas, C., Rusinkiewicz, S., Weyrich, T., 2008. A system for high-volume acquisition and matching of fresco fragments: Reassembling theran wall paintings, in: ACM transactions on graphics (TOG), ACM. p. 84.

Bushberg, J.T., Boone, J.M., 2011. The essential physics of medical imaging. Lippincott Williams & Wilkins.

Bustamante, M., Petersson, S., Eriksson, J., Alehagen, U., Dyverfeldt, P., Carlhäll, C.J., Ebbers, T., 2015. Atlas-based analysis of 4d flow cmr: automated vessel segmentation and flow quantification. Journal of Cardiovascular Magnetic Resonance 17, 87.

Cabezas, M., Oliver, A., Lladó, X., Freixenet, J., Cuadra, M.B., 2011. A review of atlas-based segmentation for magnetic resonance brain images. Computer methods and programs in biomedicine 104, e158–e177.

Cao, X., Yang, J., Zhang, J., Nie, D., Kim, M., Wang, Q., Shen, D., 2017. Deformable image registration based on similarity-steered cnn regression, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 300–308.

Carvalho, A.L., Nishimoto, I.N., Califano, J.A., Kowalski, L.P., 2005. Trends in incidence and prognosis for head and neck cancer in the United States: a site-specific analysis of the SEER database. International journal of cancer 114, 806–816.

Castillo, R., Castillo, E., Guerra, R., Johnson, V.E., McPhail, T., Garg, A.K., Guerrero, T., 2009. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. Physics in medicine and biology 54, 1849. doi:10.1002/mrm.25634.

Ceranka, J., Polfliet, M., Lecouvet, F., Michoux, N., de Mey, J., Vandemeulebroucke, J., 2018. Registration strategies for multi-modal whole-body mri mosaicing. Magnetic resonance in medicine 79, 1684–1695.

Chang, H., Fitzpatrick, J.M., 1992. A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities. IEEE Transactions on medical imaging 11, 319–329.

Charlton, R.A., Schiavone, F., Barrick, T., Morris, R., Markus, H., 2010. Diffusion tensor imaging detects age related white matter change over a 2 year follow-up which is associated with working memory decline. Journal of Neurology, Neurosurgery & Psychiatry 81, 13–19.

Cheng, K., Wei, C., Chang, S., 2004. Locating landslides using multi-temporal satellite images. Advances in Space Research 33, 296–301.

Christensen, G.E., Geng, X., Kuhl, J.G., Bruss, J., Grabowski, T.J., Pirwani, I.A., Vannier, M.W., Allen, J.S., Damasio, H., 2006. Introduction to the non-rigid image registration evaluation project (NIREP), in: International Workshop on Biomedical Image Registration, Springer. pp. 128–135. doi:10.1007/11784012_16.

Cobbold, R.S., 2006. Foundations of biomedical ultrasound. Oxford university press.

Cocosco, C.A., Kollokian, V., Kwan, R.K.S., Pike, G.B., Evans, A.C., 1997. Brainweb: Online interface to a 3D MRI simulated brain database, in: NeuroImage, Citeseer.

Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., Marchal, G., 1995. Automated multi-modality image registration based on information theory, in: Information processing in medical imaging, pp. 263–274.

Collins, D.L., Zijdenbos, A.P., Kollokian, V., Sled, J.G., Kabani, N.J., Holmes, C.J., Evans, A.C., 1998. Design and construction of a realistic digital brain phantom. Medical Imaging, IEEE transactions on 17, 463–468.

Coolen, B.F., Poot, D.H., Liem, M.I., Smits, L.P., Gao, S., Kotek, G., Klein, S., Nederveen, A.J., 2015. Three-dimensional quantitative T1 and T2 mapping of the carotid artery: Sequence design and in vivo feasibility. Magnetic Resonance in Medicine doi:10.1002/mrm.25634.

Cootes, T.F., Edwards, G.J., Taylor, C.J., 2001. Active appearance models. IEEE Transactions on pattern analysis and machine intelligence 23, 681–685.

Cootes, T.F., Marsland, S., Twining, C.J., Smith, K., Taylor, C.J., 2004. Groupwise diffeomorphic non-rigid registration for automatic model building, in: European conference on computer vision, Springer. pp. 316–327.

Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J., 1995. Active shape models-their training and application. Computer vision and image understanding 61, 38–59. doi:10.1006/cviu.1995.1004.

Cornelius, C.P., Giessler, G.A., Wilde, F., Metzger, M.C., Mast, G., Probst, F.A., 2016. Iterations of computer-and template assisted mandibular or maxillary reconstruction with free flaps containing the lateral scapular border–Evolution of a biplanar plug-on cutting guide. Journal of Cranio-Maxillofacial Surgery 44, 229–241.

Culnane, C., Rubinstein, B.I., Teague, V., 2017. Health data in an open world. arXiv preprint arXiv:1712.05627 .

Davis, B.C., Fletcher, P.T., Bullitt, E., Joshi, S., 2010. Population shape regression from random design data. International journal of computer vision 90, 255–266.

Dewan, A.M., Yamaguchi, Y., 2009. Land use and land cover change in greater dhaka, bangladesh: Using remote sensing to promote sustainable urbanization. Applied geography 29, 390–401.

Dillon, J.K., Brown, C.B., McDonald, T.M., Ludwig, D.C., Clark, P.J., Leroux, B.G., Futran, N.D., 2015. How does the close surgical margin impact recurrence and survival when treating oral squamous cell carcinoma? Journal of Oral and Maxillofacial Surgery 73, 1182–1188.

Dong, Y., Dong, Y., Hu, G., Xu, Q., 2011. Three-dimensional reconstruction of extremity tumor regions by CT and MRI image data fusion for subject-specific preoperative assessment and planning. Computer Aided Surgery 16, 220–233.

Egger, J., Pfarrkirchner, B., Gsaxner, C., Lindner, L., Schmalstieg, D., Wallner, J., 2018. Fully convolutional mandible segmentation on a valid ground-truth dataset, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE. pp. 656–660.

Feng, Q., Zhou, Y., Li, X., Mei, Y., Lu, Z., Zhang, Y., Feng, Y., Liu, Y., Yang, W., Chen, W., 2016. Liver DCE-MRI registration in manifold space based on robust principal component analysis. Scientific Reports 6.

Fiedler, M., 1973. Algebraic connectivity of graphs. Czechoslovak mathematical journal 23, 298–305.

Fitzpatrick, J.M., West, J.B., 2001. The distribution of target registration error in rigid-body point-based registration. IEEE transactions on medical imaging 20, 917–927.

Fletcher, P.T., Venkatasubramanian, S., Joshi, S., 2009. The geometric median on riemannian manifolds with application to robust atlas estimation. NeuroImage 45, S143–S152. doi:10.1016/j.neuroimage.2008.10.052.

Fortunati, V., Verhaart, R.F., Angeloni, F., van der Lugt, A., Niessen, W.J., Veenland, J.F., Paulides, M.M., van Walsum, T., 2014. Feasibility of multimodal deformable registration for head and neck tumor treatment planning. International Journal of Radiation Oncology* Biology* Physics 90, 85–93. doi:10.1016/j.ijrobp.2014.05.027.

Fortunati, V., Verhaart, R.F., Verduijn, G.M., van der Lugt, A., Angeloni, F., Niessen, W.J., Veenland, J.F., Paulides, M.M., van Walsum, T., 2015. MRI integration into treatment planning of head and neck tumors: Can patient immobilization be avoided? Radiotherapy and Oncology 115, 191–194. doi:10.1016/j.radonc.2015.03.021.

Foskey, M., Davis, B., Goyal, L., Chang, S., Chaney, E., Strehl, N., Tomei, S., Rosenman, J., Joshi, S., 2005. Large deformation three-dimensional image registration in image-guided radiation therapy. Physics in Medicine & Biology 50, 5869.

Galbán, C.J., Han, M.K., Boes, J.L., Chughtai, K.A., Meyer, C.R., Johnson, T.D., Galbán, S., Rehemtulla, A., Kazerooni, E.A., Martinez, F.J., et al., 2012. Computed tomography–based biomarker provides unique signature for diagnosis of copd phenotypes and disease progression. Nature medicine 18, 1711.

Geng, X., Christensen, G.E., Gu, H., Ross, T.J., Yang, Y., 2009. Implicit reference-based group-wise image registration and its application to structural and functional MRI. Neuroimage 47, 1341–1351. doi:10.1016/j.neuroimage.2009.04.024.

Gerber, N., Bell, B., Gavaghan, K., Weisstanner, C., Caversaccio, M., Weber, S., 2014. Surgical planning tool for robotically assisted hearing aid implantation. International journal of computer assisted radiology and surgery 9, 11–20.

Giesel, F., Mehndiratta, A., Locklin, J., McAuliffe, M., White, S., Choyke, P., Knopp, M., Wood, B., Haberkorn, U., von Tengg-Kobligk, H., 2009. Image fusion using CT, MRI and PET for treatment planning, navigation and follow up in percutaneous RFA. Experimental oncology 31, 106.

Giles, S.L., Messiou, C., Collins, D.J., Morgan, V.A., Simpkin, C.J., West, S., Davies, F.E., Morgan, G.J., deSouza, N.M., 2014. Whole-body diffusion-weighted mr imaging for assessment of treatment response in myeloma. Radiology 271, 785–794.

Golparvar-Fard, M., Peña-Mora, F., Savarese, S., 2009. D4ar–a 4-dimensional augmented reality model for automating construction progress monitoring data collection, processing and communication. Journal of information technology in construction 14, 129–153.

Gorbunova, V., Lo, P., Ashraf, H., Dirksen, A., Nielsen, M., de Bruijne, M., 2008. Weight preserving image registration for monitoring disease progression in lung ct, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 863–870.

Guerrero, R., Wolz, R., Rueckert, D., 2011. Laplacian eigenmaps manifold learning for landmark localization in brain MR images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 566–573.

Guyader, J.M., Bernardin, L., Douglas, N.H., Poot, D.H., Niessen, W.J., Klein, S., 2015. Influence of image registration on apparent diffusion coefficient images computed from free-breathing diffusion mr images of the abdomen. Journal of Magnetic Resonance Imaging 42, 315–330.

Guyader, J.M., Huizinga, W., Poot, D.H., van Kranenburg, M., Uitterdijk, A., Niessen, W.J., Klein, S., 2018. Groupwise image registration based on a total correlation dissimilarity measure for quantitative MRI and dynamic imaging data. Scientific reports 8, 13112.

Hallack, A., Chappell, M.A., Gooding, M.J., Schnabel, J.A., 2014. A new similarity metric for groupwise registration of variable flip angle sequences for improved T 10 estimation in DCE-MRI, in: Biomedical Image Registration. Springer, pp. 154–163. doi:10.1007/978-3-319-08554-8_16.

Hamm, J., Davatzikos, C., Verma, R., 2009. Efficient large deformation registration via geodesics on a learned manifold of images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 680–687.

Hamrouni, S., Rougon, N., Prêteux, F., 2011. Multi-feature information-theoretic image registration: application to groupwise registration of perfusion mri exams, in: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE. pp. 574–577.

Hamy, V., Dikaios, N., Punwani, S., Melbourne, A., Latifoltojar, A., Makanyanga, J., Chouhan, M., Helbren, E., Menys, A., Taylor, S., et al.,

2014. Respiratory motion correction in dynamic MRI using robust data decomposition registration–application to DCE-MRI. Medical image analysis 18, 301–313.

Hero, A., Ma, B., Michel, O.J., Gorman, J., 2002. Applications of entropic spanning graphs. Signal Processing Magazine, IEEE 19, 85–95. doi:10.1109/msp.2002.1028355.

Hsieh, J., et al., 2009. Computed tomography: principles, design, artifacts, and recent advances, SPIE Bellingham, WA.

Huang, X., Hill, N.A., Ren, J., Guiraudon, G., Boughner, D., Peters, T.M., 2005. Dynamic 3d ultrasound and mr image registration of the beating heart, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 171–178.

Huizinga, W., Poot, D., Guyader, J.M., Klaassen, R., Coolen, B., van Kranenburg, M., van Geuns, R., Uitterdijk, A., Polfliet, M., Vandemeulebroucke, J., Leemans, A., Niessen, W., Klein, S., 2016. PCA-based groupwise image registration for quantitative MRI. Medical Image Analysis 29, 65–78. doi:10.1016/j.media.2015.12.004.

Huizinga, W., Poot, D.H., Vernooij, M.W., Roshchupkin, G., Bron, E., Ikram, M.A., Rueckert, D., Niessen, W.J., Klein, S., Initiative, A.D.N., et al., 2018. A spatio-temporal reference model of the aging brain. NeuroImage 169, 11–22.

Irani, M., Peleg, S., 1991. Improving resolution by image registration. CVGIP: Graphical models and image processing 53, 231–239.

Ireland, R.H., Dyker, K.E., Barber, D.C., Wood, S.M., Hanney, M.B., Tindale, W.B., Woodhouse, N., Hoggard, N., Conway, J., Robinson, M.H., 2007. Nonrigid image registration for head and neck cancer radiotherapy treatment planning with pet/ct. International Journal of Radiation Oncology* Biology* Physics 68, 952–957.

Isgum, I., Staring, M., Rutten, A., Prokop, M., Viergever, M.A., van Ginneken, B., 2009. Multi-atlas-based segmentation with local decision fusionapplication to cardiac and aortic segmentation in ct scans. IEEE transactions on medical imaging 28, 1000–1010.

James, A.P., Dasarathy, B.V., 2014. Medical image fusion: A survey of the state of the art. Information Fusion 19, 4–19.

Joshi, S., Davis, B., Jomier, M., Gerig, G., 2004. Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage 23, S151–S160. doi:10.1016/j.neuroimage.2004.07.068.

Juerchott, A., Freudlsperger, C., Zingler, S., Saleem, M.A., Jende, J.M., Lux, C.J., Bendszus, M., Heiland, S., Hilgenfeld, T., 2020. In vivo reliability of 3d cephalometric landmark determination on magnetic resonance imaging: A feasibility study. Clinical Oral Investigations 24, 1339–1349.

Kademani, D., Bell, R.B., Bagheri, S., Holmgren, E., Dierks, E., Potter, B., Homer, L., 2005. Prognostic factors in intraoral squamous cell carcinoma: the influence of histologic grade. Journal of oral and maxillofacial surgery 63, 1599–1605.

Kim, M., Wu, G., Shen, D., 2011. Groupwise registration of breast DCE-MR images for accurate tumor measurement, in: Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, IEEE. pp. 598–601.

Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., et al., 2009a. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 46, 786–802.

Klein, S., Pluim, J.P., Staring, M., Viergever, M.A., 2009b. Adaptive stochastic gradient descent optimisation for image registration. International journal of computer vision 81, 227–239. doi:10.1007/s11263-008-0168-y.

Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P., 2010. Elastix: a toolbox for intensity-based medical image registration. Medical Imaging, IEEE Transactions on 29, 196–205. doi:10.1109/tmi.2009.2035616.

Klein, S., Staring, M., Pluim, J.P., 2007. Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. IEEE transactions on image processing 16, 2879–2890.

van 't Klooster, R., Staring, M., Klein, S., Kwee, R.M., Kooi, M.E., Reiber, J.H., Lelieveldt, B.P., van der Geest, R.J., 2013. Automated registration of multispectral mr vessel wall images of the carotid artery. Medical physics 40.

König, L., Derksen, A., Papenberg, N., Haas, B., 2016. Deformable image registration for adaptive radiotherapy with guaranteed local rigidity constraints. Radiation Oncology 11, 122.

Kraeima, J., Dorgelo, B., Gulbitti, H., Steenbakkers, R., Schepman, K., Roodenburg, J., Spijkervet, F., Schepers, R., Witjes, M., 2018. Multi-modality 3d mandibular resection planning in head and neck cancer using ct and mri data fusion: A clinical series. Oral oncology 81, 22–28.

Kraeima, J., Schepers, R.H., van Ooijen, P.M., Steenbakkers, R.J., Roodenburg, J.L., Witjes, M.J., 2015. Integration of oncologic margins in three-dimensional virtual planning for head and neck surgery, including a validation of the software pathway. Journal of Cranio-Maxillofacial Surgery 43, 1374–1379.

Kreppel, M., Drebber, U., Eich, H.T., Dreiseidler, T., Zöller, J.E., Müller, R.P., Scheer, M., 2011. Combined-modality treatment in advanced oral squamous cell carcinoma. Strahlentherapie und Onkologie 187, 555–560.

Krish, R.P., Fierrez, J., Ramos, D., Ortega-Garcia, J., Bigun, J., 2015. Preregistration of latent fingerprints based on orientation field. IET Biometrics 4, 42–52.

Lagravère, M.O., Low, C., Flores-Mir, C., Chung, R., Carey, J.P., Heo, G., Major, P.W., 2010. Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images. American Journal of Orthodontics and Dentofacial Orthopedics 137, 598–604.

Lanczos, C., 1950. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. United States Governm. Press Office Los Angeles, CA.

Learned-Miller, E.G., 2006. Data driven image models through continuous joint alignment. Pattern Analysis and Machine Intelligence, IEEE Transactions on 28, 236–250. doi:`10.1109/tpami.2006.34`.

de Leeuw, J., 2007. Derivatives of generalized eigen systems with applications. Department of Statistics, UCLA .

Lei, Y., Fu, Y., Wang, T., Liu, Y., Patel, P., Curran, W.J., Liu, T., Yang, X., 2020. 4d-ct deformable image registration using multiscale unsupervised deep learning. Physics in Medicine & Biology 65, 085003.

Leibfarth, S., Mönnich, D., Welz, S., Siegel, C., Schwenzer, N., Schmidt, H., Zips, D., Thorwarth, D., 2013. A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning. Acta oncologica 52, 1353–1359.

Li, C., Men, Y., Yang, W., Pan, J., Sun, J., Li, L., 2014a. Computed tomography for the diagnosis of mandibular invasion caused by head and neck cancer: a systematic review comparing contrast-enhanced and plain computed tomography. Journal of Oral and Maxillofacial Surgery 72, 1601–1615.

Li, C., Yang, W., Men, Y., Wu, F., Pan, J., Li, L., 2014b. Magnetic resonance imaging for diagnosis of mandibular involvement from head and neck cancers: a systematic review and meta-analysis. PloS one 9, e112267.

Lombaert, H., Grady, L., Pennec, X., Ayache, N., Cheriet, F., 2014. Spectral log-demons: diffeomorphic image registration with very large deformations. International journal of computer vision 107, 254–271.

Lorenzi, M., Ayache, N., Frisoni, G.B., Pennec, X., (ADNI, A.D.N.I., et al., 2013. Lcc-demons: a robust and accurate symmetric diffeomorphic registration algorithm. NeuroImage 81, 470–483.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. International journal of computer vision 60, 91–110.

Lu, W., Olivera, G.H., Chen, Q., Ruchala, K.J., Haimerl, J., Meeks, S.L., Langen, K.M., Kupelian, P.A., 2006. Deformable registration of the planning image (kvct) and the daily images (mvct) for adaptive radiation therapy. Physics in Medicine & Biology 51, 4357.

Ludlow, J.B., Gubler, M., Cevidanes, L., Mol, A., 2009. Precision of cephalometric landmark identification: cone-beam computed tomography vs conventional cephalometric views. American Journal of Orthodontics and Dentofacial Orthopedics 136, 312–e1.

Luenberger, D.G., 1973. Introduction to linear and nonlinear programming. volume 28. Addison-Wesley Reading, MA.

Lynch, J.A., Zaim, S., Zhao, J., Peterfy, C.G., Genant, H.K., 2001. Automating measurement of subtle changes in articular cartilage from mri of the knee by combining 3d image registration and segmentation, in: Medical Imaging 2001: Image Processing, International Society for Optics and Photonics. pp. 431–439.

Ma, J., Miller, M.I., Trouvé, A., Younes, L., 2008. Bayesian template estimation in computational anatomy. NeuroImage 42, 252–261.

Mackie, T.R., Kapatoes, J., Ruchala, K., Lu, W., Wu, C., Olivera, G., Forrest, L., Tome, W., Welsh, J., Jeraj, R., et al., 2003. Image guidance for precise conformal radiotherapy. International Journal of Radiation Oncology* Biology* Physics 56, 89–105.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997. Multimodality image registration by maximization of mutual information. Medical Imaging, IEEE Transactions on 16, 187–198. doi:10.1109/42.563664.

Makropoulos, A., Counsell, S.J., Rueckert, D., 2017. A review on automatic fetal and neonatal brain MRI segmentation. NeuroImage .

Melbourne, A., Atkinson, D., White, M., Collins, D., Leach, M., Hawkes, D., 2007. Registration of dynamic contrast-enhanced MRI using a progressive principal component registration (PPCR). Physics in medicine and biology 52, 5147.

Metz, C., Klein, S., Schaap, M., van Walsum, T., Niessen, W.J., 2011. Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach. Medical image analysis 15, 238–249. doi:10.1016/j.media.2010.10.003.

Michopoulou, S.K., Costaridou, L., Panagiotopoulos, E., Speller, R., Panayiotakis, G., Todd-Pokropek, A., 2009. Atlas-based segmentation of degenerated lumbar intervertebral discs from mr images of the spine. IEEE transactions on Biomedical Engineering 56, 2225–2231.

Miller, M.I., Priebe, C.E., Qiu, A., Fischl, B., Kolasny, A., Brown, T., Park, Y., Ratnanather, J.T., Busa, E., Jovicich, J., et al., 2009. Collaborative computational anatomy: an mri morphometry study of the human brain via diffeomorphic metric mapping. Human brain mapping 30, 2132–2141.

Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S., 2010. Fast free-form deformation using graphics processing units. Computer methods and programs in biomedicine 98, 278–284.

Mohar, B., Alavi, Y., Chartrand, G., Oellermann, O., 1991. The laplacian spectrum of graphs. Graph theory, combinatorics, and applications 2, 12.

Murphy, K., van Ginneken, B., Klein, S., Staring, M., de Hoop, B.J., Viergever, M.A., Pluim, J.P., 2011a. Semi-automatic construction of reference standards for evaluation of image registration. Medical image analysis 15, 71–84.

Murphy, K., van Ginneken, B., Reinhardt, J.M., Kabus, S., Ding, K., Deng, X., Cao, K., Du, K., Christensen, G.E., Garcia, V., et al., 2011b. Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge. IEEE transactions on medical imaging 30, 1901–1920.

Neemuchwala, H., Hero, A., Zabuawala, S., Carson, P., 2006. Image registration methods in high-dimensional space. International Journal of Imaging Systems and Technology 16, 130–145.

Nemec, S.F., Peloschek, P., Schmook, M.T., Krestan, C.R., Hauff, W., Matula, C., Czerny, C., 2010. Ct–mr image data fusion for computer-assisted navigated surgery of orbital tumors. European journal of radiology 73, 224–229.

Ng, C.S., Chandler, A.G., Wei, W., Herron, D.H., Anderson, E.F., Kurzrock, R., Charnsangavej, C., 2011. Reproducibility of ct perfusion parameters in liver tumors and normal liver. Radiology 260, 762–770.

Nieberler, M., Häußler, P., Kesting, M.R., Kolk, A., Deppe, H., Weirich, G., Wolff, K.D., 2016. Clinical impact of intraoperative cytological assessment of bone resection margins in patients with head and neck carcinoma. Annals of surgical oncology 23, 3579–3586.

Nishioka, T., Shiga, T., Shirato, H., Tsukamoto, E., Tsuchiya, K., Kato, T., Ohmori, K., Yamazaki, A., Aoyama, H., Hashimoto, S., et al., 2002. Image fusion between 18fdg-pet and mri/ct for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. International Journal of Radiation Oncology* Biology* Physics 53, 1051–1057.

Nishiyama, K.K., Pauchard, Y., Nikkel, L.E., Iyer, S., Zhang, C., McMahon, D.J., Cohen, D., Boyd, S.K., Shane, E., Nickolas, T.L., 2015. Longitudinal hr-pqct and image registration detects endocortical bone loss in kidney transplantation patients. Journal of Bone and Mineral Research 30, 554–561.

Oh, S., Kim, S., 2017. Deformable image registration in radiation therapy. Radiation oncology journal 35, 101.

Oliveira, F.P., Tavares, J.M.R., 2014. Medical image registration: a review. Computer methods in biomechanics and biomedical engineering 17, 73–93.

Orchard, J., Mann, R., 2010. Registering a multisensor ensemble of images. Image Processing, IEEE Transactions on 19, 1236–1247. doi:10.1109/tip.2009.2039371.

Pennec, X., Cachier, P., Ayache, N., 2003. Tracking brain deformations in time sequences of 3d us images. Pattern Recognition Letters 24, 801–813.

Piella, G., 2014. Diffusion maps for multimodal registration. Sensors 14, 10562–10577.

Pluim, J.P., Maintz, J.A., Viergever, M.A., 2003. Mutual-information-based registration of medical images: a survey. Medical Imaging, IEEE Transactions on 22, 986–1004. doi:`10.1109/tmi.2003.815867`.

Polfliet, M., Klein, S., Huizinga, W., de Mey, J., Vandemeulebroucke, J., 2016. The Pythagorean averages as group images in efficient groupwise registration, in: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1261–1264. doi:`10.1109/isbi.2016.7493496`.

Polfliet, M., Klein, S., Huizinga, W., Paulides, M.M., Niessen, W.J., Vandemeulebroucke, J., 2018. Intrasubject multimodal groupwise registration with the conditional template entropy. Medical image analysis 46, 15–25.

Polfliet, M., Klein, S., Niessen, W.J., Vandemeulebroucke, J., 2017. Laplacian eigenmaps for multimodal groupwise image registration, in: SPIE Medical Imaging, International Society for Optics and Photonics. pp. 101331N–101331N.

Reuter, M., Rosas, H.D., Fischl, B., 2010. Highly accurate inverse consistent registration: a robust approach. Neuroimage 53, 1181–1196.

Risholm, P., Golby, A.J., Wells, W., 2011. Multimodal image registration for preoperative planning and image-guided neurosurgical procedures. Neurosurgery Clinics 22, 197–206.

Ristic, M., Brujic, D., 1997. Efficient registration of nurbs geometry. Image and vision computing 15, 925–935.

Rivaz, H., Chen, S.J.S., Collins, D.L., 2015. Automatic deformable MR-ultrasound registration for image-guided neurosurgery. IEEE transactions on medical imaging 34, 366–380.

Rivest-Hénault, D., Dowson, N., Greer, P.B., Fripp, J., Dowling, J.A., 2015. Robust inverse-consistent affine CT–MR registration in MRI-assisted and MRI-alone prostate radiation therapy. Medical image analysis 23, 56–69.

Roche, A., Malandain, G., Ayache, N., Pennec, X., et al., 1998a. Multimodal image registration by maximization of the correlation ratio .

Roche, A., Malandain, G., Pennec, X., Ayache, N., 1998b. The correlation ratio as a new similarity measure for multimodal image registration, in: Medical Image Computing and Computer-Assisted Interventation MICCAI98. Springer, pp. 1115–1124. doi:10.1007/bfb0056301.

Roche, A., Pennec, X., Malandain, G., Ayache, N., 2001. Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information. IEEE transactions on medical imaging 20, 1038–1049.

Rohlfing, T., 2011. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. IEEE transactions on medical imaging 31, 153–163.

Roweis, S.T., Saul, L.K., 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 290, 2323–2326.

Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. Medical Imaging, IEEE Transactions on 18, 712–721. doi:10.1109/42.796284.

Safi, A.F., Kauke, M., Grandoch, A., Nickenig, H.J., Zöller, J.E., Kreppel, M., 2017. Analysis of clinicopathological risk factors for locoregional recurrence of oral squamous cell carcinoma–Retrospective analysis of 517 patients. Journal of Cranio-Maxillofacial Surgery 45, 1749–1753.

Saito, T., Sakamoto, T., Oya, N., 2009. Comparison of gating around end-expiration and end-inspiration in radiotherapy for lung cancer. Radiotherapy and oncology 93, 430–435.

Sanz-Estébanez, S., Pena-Nogales, O., de Luis-García, R., Aja-Fernández, S., Alberola-López, C., 2017. Groupwise non-rigid registration on multiparametric abdominal dwi acquisitions for robust adc estimation: Comparison with pairwise approaches and different multimodal metrics, in: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), IEEE. pp. 1156–1159.

Sarrut, D., Boldea, V., Miguet, S., Ginestet, C., 2006. Simulation of four-dimensional CT images from deformable registration between inhale and exhale breath-hold CT scans. Medical physics 33, 605–617. doi:10.1118/1.2161409.

Sedghi, A., Luo, J., Mehrtash, A., Pieper, S., Tempany, C.M., Kapur, T., Mousavi, P., Wells III, W.M., 2018. Semi-supervised deep metrics for image registration. arXiv preprint arXiv:1804.01565 .

Serag, A., Aljabar, P., Ball, G., Counsell, S.J., Boardman, J.P., Rutherford, M.A., Edwards, A.D., Hajnal, J.V., Rueckert, D., 2012. Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression. Neuroimage 59, 2255–2265.

Shah, J.P., Gil, Z., 2009. Current concepts in management of oral cancer–surgery. Oral oncology 45, 394–401.

Shalaby, A., Tateishi, R., 2007. Remote sensing and gis for mapping and monitoring land cover and land-use changes in the northwestern coastal zone of egypt. Applied Geography 27, 28–41.

Sharma, A., Horaud, R.P., Knossow, D., Von Lavante, E., 2009. Mesh segmentation using Laplacian eigenvectors and Gaussian mixtures., in: AAAI Fall Symposium: Manifold Learning and Its Applications.

Sidorov, K.A., 2018. Estimating the age of birch bark manuscripts using computational paleography. Programming and Computer Software 44, 286–294.

Simone, G., Farina, A., Morabito, F.C., Serpico, S.B., Bruzzone, L., 2002. Image fusion techniques for remote sensing applications. Information fusion 3, 3–15.

Skerl, D., Likar, B., Pernus, F., 2006. A protocol for evaluation of similarity measures for rigid registration. Medical Imaging, IEEE Transactions on 25, 779–791. doi:10.1109/tmi.2006.874963.

Smits, R., Ten Hove, I., Dronkers, E., Schut, T.B., Mast, H., de Jong, R.B., Wolvius, E., Puppels, G., Koljenović, S., 2018. Evaluation of bone resection margins of segmental mandibulectomy for oral squamous cell carcinoma. International journal of oral and maxillofacial surgery 47, 959–964.

Sokooti, H., de Vos, B., Berendsen, F., Lelieveldt, B.P., Išgum, I., Staring, M., 2017. Nonrigid image registration using multi-scale 3d convolutional neural

networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 232–239.

Sotiras, A., Davatzikos, C., Paragios, N., 2013. Deformable medical image registration: A survey. IEEE transactions on medical imaging 32, 1153–1190.

Spiclin, Z., Likar, B., Pernus, F., 2012. Groupwise registration of multimodal images by an efficient joint entropy minimization scheme. Image Processing, IEEE Transactions on 21, 2546–2558. doi:10.1109/tip.2012.2186145.

Staring, M., Klein, S., Pluim, J.P., 2007. A rigidity penalty term for nonrigid registration. Medical physics 34, 4098–4108.

Studholme, C., Cardenas, V., 2004. A template free approach to volumetric spatial normalization of brain anatomy. Pattern Recognition Letters 25, 1191–1202.

Studholme, C., Hill, D.L., Hawkes, D.J., 1999. An overlap invariant entropy measure of 3D medical image alignment. Pattern recognition 32, 71–86. doi:10.1016/s0031-3203(98)00091-0.

Succo, G., Berrone, M., Battiston, B., Tos, P., Goia, F., Appendino, P., Crosetti, E., 2015. Step-by-step surgical technique for mandibular reconstruction with fibular free flap: application of digital technology in virtual surgical planning. European Archives of Oto-Rhino-Laryngology 272, 1491–1501.

Sun, W., Niessen, W.J., Klein, S., 2017. Randomly perturbed B-splines for non-rigid image registration. IEEE transactions on pattern analysis and machine intelligence 39, 1401–1413.

Szeliski, R., et al., 2007. Image alignment and stitching: A tutorial. Foundations and Trends® in Computer Graphics and Vision 2, 1–104.

Tang, S., Fan, Y., Wu, G., Kim, M., Shen, D., 2009. RABBIT: rapid alignment of brains by building intermediate templates. NeuroImage 47, 1277–1287.

Tarsitano, A., Ciocca, L., Cipriani, R., Scotti, R., Marchetti, C., 2015. Mandibular reconstruction using fibula free flap harvested using a customised cutting guide: how we do it. Acta Otorhinolaryngologica Italica 35, 198.

Tenenbaum, J.B., De Silva, V., Langford, J.C., 2000. A global geometric framework for nonlinear dimensionality reduction. science 290, 2319–2323.

Thévenaz, P., Unser, M., 2000. Optimization of mutual information for multiresolution image registration. Image Processing, IEEE Transactions on 9, 2083–2099. doi:10.1109/83.887976.

Thirion, J.P., 1998. Image matching as a diffusion process: an analogy with maxwell's demons. Medical image analysis 2, 243–260.

Tomaževič, D., Likar, B., Pernuš, F., 2012. Multi-feature mutual information image registration. Image Analysis & Stereology 31, 43–53. doi:10.5566/ias.v31.p43-53.

Royuela-del Val, J., Cordero-Grande, L., Simmross-Wattenberg, F., Martín-Fernández, M., Alberola-López, C., 2016. Nonrigid groupwise registration for motion estimation and compensation in compressed sensing reconstruction of breath-hold cardiac cine MRI. Magnetic resonance in medicine 75, 1525–1536.

Van Assche, N., Van Steenberghe, D., Guerrero, M., Hirsch, E., Schutyser, F., Quirynen, M., Jacobs, R., 2007. Accuracy of implant placement based on pre-surgical planning of three-dimensional cone-beam images: a pilot study. Journal of clinical periodontology 34, 816–821.

Vandemeulebroucke, J., Bernard, O., Rit, S., Kybic, J., Clarysse, P., Sarrut, D., 2012. Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT. Medical physics 39, 1006–1015. doi:10.1118/1.3679009.

Vandemeulebroucke, J., Rit, S., Kybic, J., Clarysse, P., Sarrut, D., 2011. Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. Medical physics 38, 166–178. doi:10.1118/1.3523619.

Varvares, M.A., Poti, S., Kenyon, B., Christopher, K., Walker, R.J., 2015. Surgical margins and primary site resection in achieving local control in oral cancer resections. The Laryngoscope 125, 2298–2307.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2009. Diffeomorphic demons: Efficient non-parametric image registration. NeuroImage 45, S61–S72.

Verhaart, R.F., Fortunati, V., Verduijn, G.M., van der Lugt, A., van Walsum, T., Veenland, J.F., Paulides, M.M., 2014. The relevance of MRI for patient modeling in head and neck hyperthermia treatment planning: A comparison of CT and CT-MRI based tissue segmentation on simulated temperature. Medical physics 41, 123302. doi:10.1118/1.4901270.

Vijayan, S., Klein, S., Hofstad, E.F., Lindseth, F., Ystgaard, B., Langø, T., 2014. Motion tracking in the liver: Validation of a method based on 4D ultrasound using a nonrigid registration technique. Medical physics 41. doi:10.1118/1.4890091.

Viola, P., Wells III, W., 1995. Alignment by maximization of mutual information, in: Proceedings of the Fifth International Conference on Computer Vision, IEEE Computer Society. p. 16.

Vishnevskiy, V., Gass, T., Szekely, G., Tanner, C., Goksel, O., 2016. Isotropic total variation regularization of displacements in parametric image registration. IEEE transactions on medical imaging 36, 385–395.

de Vos, B.D., Berendsen, F.F., Viergever, M.A., Sokooti, H., Staring, M., Išgum, I., 2019. A deep learning framework for unsupervised affine and deformable image registration. Medical image analysis 52, 128–143.

Wachinger, C., Glocker, B., Zeltner, J., Paragios, N., Komodakis, N., Hansen, M.S., Navab, N., 2008. Deformable mosaicing for whole-body mri, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 113–121.

Wachinger, C., Navab, N., 2012. Entropy and laplacian images: Structural representations for multi-modal registration. Medical Image Analysis 16, 1–17. doi:10.1016/j.media.2011.03.001.

Wachinger, C., Navab, N., 2013. Simultaneous registration of multiple images: Similarity metrics and efficient optimization. Pattern Analysis and Machine Intelligence, IEEE Transactions on 35, 1221–1233. doi:10.1109/tpami.2012.196.

Wachinger, C., Wein, W., Navab, N., 2007. Three-dimensional ultrasound mosaicing, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007. Springer, pp. 327–335. doi:10.1007/978-3-540-75759-7_40.

Wallner, J., Schwaiger, M., Hochegger, K., Gsaxner, C., Zemann, W., Egger, J., 2019. A review on multiplatform evaluations of semi-automatic open-source based image segmentation for cranio-maxillofacial surgery. Computer methods and programs in biomedicine 182, 105102.

Wang, L., Nie, D., Li, G., Puybareau, É., Dolz, J., Zhang, Q., Wang, F., Xia, J., Wu, Z., Chen, J., et al., 2019. Benchmark on automatic 6-month-old infant brain segmentation algorithms: The iSeg-2017 challenge. IEEE transactions on medical imaging .

Wang, Z., Ma, Y., 2008. Medical image fusion using m-pcnn. Information Fusion 9, 176–185.

Warfield, S.K., Zou, K.H., Wells, W.M., 2002. Validation of image segmentation and expert quality with an expectation-maximization algorithm, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 298–306.

Webster, G.J., Kilgallon, J., Ho, K.F., Rowbottom, C.G., Slevin, N.J., Mackay, R.I., 2009. A novel imaging technique for fusion of high-quality immobilised MR images of the head and neck with CT scans for radiotherapy target delineation. The British Journal of Radiology 82, 497–503.

Wells, J.A., O'Callaghan, J., Holmes, H., Powell, N.M., Johnson, R., Siow, B., Torrealdea, F., Ismail, O., Walker-Samuel, S., Golay, X., et al., 2015. In vivo imaging of tau pathology using multi-parametric quantitative MRI. Neuroimage 111, 369–378. doi:10.1016/j.neuroimage.2015.02.023.

Wells, W.M., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R., 1996. Multimodal volume registration by maximization of mutual information. Medical image analysis 1, 35–51. doi:10.1016/s1361-8415(96)80004-1.

West, J., Fitzpatrick, J.M., Wang, M.Y., Dawant, B.M., Maurer Jr, C.R., Kessler, R.M., Maciunas, R.J., Barillot, C., Lemoine, D., Collignon, A., Maes, F., Suetens, P., Vandermeulen, D., van den Elsen, P.A., Napel, S., Sumanaweera, T.S., Harkness, B., Hemler, P.F., Hill, D.L., Hawkes, D.J., Studholme, C., Maintz, J.A., Viergever, M.A., Malandain, G., Pennec, X., Noz, M.E., Maguire, G.Q.J., Pollack, M., Pelizzari, C.A., Robb, R.A., Hanson, D., Woods, R.P., 1997. Comparison and evaluation of retrospective intermodality brain image registration techniques. Journal of computer assisted tomography 21, 554–568. doi:10.1097/00004728-199707000-00007.

Williams, F.L., Richtsmeier, J.T., 2003. Comparison of mandibular landmarks from computed tomography and 3d digitizer data. Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists 16, 494–500.

Williams, T., Holmes, A., Bowes, M., Vincent, G., Hutchinson, C., Waterton, J., Maciewicz, R., Taylor, C., 2010. Measurement and visualisation of focal cartilage thickness change by MRI in a study of knee osteoarthritis using a novel image analysis tool. The British journal of radiology 83, 940–948.

Wolthaus, J.W., Schneider, C., Sonke, J.J., van Herk, M., Belderbos, J.S., Rossi, M.M., Lebesque, J.V., Damen, E.M., 2006. Mid-ventilation CT scan construction from four-dimensional respiration-correlated CT scans for radiotherapy planning of lung cancer patients. International Journal of Radiation Oncology* Biology* Physics 65, 1560–1571.

Wu, G., Jia, H., Wang, Q., Shen, D., 2011. Sharpmean: Groupwise registration guided by sharp mean image and tree-based registration. NeuroImage 56, 1968–1981.

Wu, G., Peng, X., Ying, S., Wang, Q., Yap, P.T., Shen, D., Shen, D., 2016. eHUGS: enhanced hierarchical unbiased graph shrinkage for efficient groupwise registration. PloS one 11, e0146870.

Xue, H., Srinivasan, L., Jiang, S., Rutherford, M., Edwards, A.D., Rueckert, D., Hajnal, J.V., 2007. Automatic segmentation and reconstruction of the cortex from neonatal MRI. Neuroimage 38, 461–477.

Yigitsoy, M., Wachinger, C., Navab, N., 2011. Temporal groupwise registration for motion modeling, in: Information Processing in Medical Imaging, Springer. pp. 648–659.

Yu, H., Jiang, H., Zhou, X., Hara, T., Yao, Y.D., Fujita, H., 2020. Unsupervised 3d pet-ct image registration method using a metabolic constraint function and a multi-domain similarity measure. IEEE Access 8, 63077–63089.

Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G., 2006. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 31, 1116–1128.

Zimmer, V.A., Ballester, M.Á.G., Piella, G., 2019. Multimodal image registration using Laplacian commutators. Information Fusion 49, 130–145.

Zitova, B., Flusser, J., 2003. Image registration methods: a survey. Image and vision computing 21, 977–1000.

Zöllei, L., Learned-Miller, E., Grimson, E., Wells, W., 2005. Efficient population registration of 3D data, in: Computer Vision for Biomedical Image Applications. Springer, pp. 291–301. doi:10.1007/11569541_30.

Zvitia, O., Mayer, A., Shadmi, R., Miron, S., Greenspan, H.K., 2010. Co-registration of white matter tractographies by adaptive-mean-shift and gaussian mixture modeling. IEEE Transactions on Medical Imaging 29, 132–145.

CHAPTER 8

PhD Portfolio

PhD period: 2013-2022
Departments: Department of electronics and informatics, VUB & Department of radiology and nuclear medicine, Erasmus MC
Research schools: Natural Sciences and (Bioscience) Engineering, VUB & Advanced School for Computing and Imaging, EUR

Courses, workshops and summer schools

Medical Imaging Summer School - 3 ECTS 2014
IWT Defense Training - 1 ECTS 2014
Managing my PhD in Natural and Applied Sciences - 1 ECTS 2015
Writing Articles the Natural and Applied Sciences - 1 ECTS 2017
Efficient networking skills - 1 ECTS 2017
Statistics for PhD candidates - 3 ECTS 2017
Research integrity - 0.3 ECTS 2018

Total - 10.3 ECTS

International conferences

Workshop on Biomedical Image Registration (Londen) - Attendance - 1 ECTS 2014
European Conference of the International Federation for Medical and Biological Engineering - Oral presentation - 2 ECTS 2014

International Conference on Medical Image Computing and Computer Assisted Intervention - Attendance - 1 ECTS 2015
Imaging and Computer Assistance in Radiation Therapy Workshop - Oral presentation - 2 ECTS 2015
International Symposium on Biomedical Imaging - Poster - 1 ECTS 2016
SPIE: Medical Imaging - Poster Presentation - 2 ECTS 2017

Total - 9 ECTS

National conferences

Superminds - Poster presentation - 1 ECTS 2015
National day of biomedical engineering - Poster presentation - 2 ECTS 2015-2016

Total - 3 ECTS

Teaching

Modeling of physiological systems - 20 ECTS 2013-2018
Biomedical imaging - 4 ECTS 2014-2015
Informatica I - 8 ECTS 2016-2018
Health information and decision support systems - 4 ECTS 2017-2018

Supervision of master thesis 'Comparison of Registration-Based Measures of Lung Ventilation from CT with Multiple-Breath Nitrogen Washout Tests' by Maëliss Avenoso - 3 ECTS 2015-2016
Supervision of master thesis 'Optimizing Lung Registration for CT Ventilation Estimation' by Isabel Lara Riera - 3 ECTS 2016-2017

Total - 42 ECTS

Others

Co-organizing 'National Day of BioMedical Engineering' - 3 ECTS 2015
Secretary for the study programme committee of the Master of Science in Biomedical engineering - 3 ECTS 2015-2018

Planning and scheduling timetables - Master of Science in Biomedical engineering (VUB) - 5 ECTS 2015-2018
Reviewing for 'Transactions on Medical Imaging' - 1 ECTS 2016

Total - 12 ECTS

CHAPTER 9

Publications

Publications in international journals

- De Pauw, K., Roelands, B., Knaepen, K., **Polfliet, M.**, Stiens, J. & Meeusen, R., 2015. Effects of caffeine and maltodextrin mouth rinsing on P300, brain imaging, and cognitive performance. Journal of Applied Physiology, vol. 118, no. 6, pp. 776-782.

- Scheerlinck, T., **Polfliet, M.**, Deklerck, R., Van Gompel, G., Buls, N. & Vandemeulebroucke, J., 2016. Development and validation of an automated and marker-free CT-based spatial analysis method (CTSA) for assessment of femoral hip implant migration In vitro accuracy and precision comparable to that of radiostereometric analysis (RSA). Acta Orthopaedica, vol. 87, no. 2, pp. 139-145.

- Huizinga, W., Poot, D.H.J., Guyader, J.-M., Klaassen, R., Coolen, B.F., van Kranenburg, M., van Geuns, R.J.M., Uitterdijk, A, **Polfliet, M.**, Vandemeulebroucke, J., Leemans, A., Niessen, W.J. & Klein, S., 2016. PCA-based groupwise image registration for quantitative MRI. Medical Image Analysis, vol. 29, pp. 65-78.

- Ceranka, J., **Polfliet, M.**, Lecouvet, F., Michoux, N., de Mey, J. & Vandemeulebroucke, J., 2018. Registration strategies for multi-modal whole-body MRI mosaicing. Magnetic Resonance in Medicine, vol. 79, no. 3, pp. 1684-1695.

- **Polfliet, M.**, Klein, S., Huizinga, W., Paulides, M.M., Niessen, W.J. & Vandemeulebroucke, J., 2018. Intrasubject multimodal groupwise regis-

tration with the conditional template entropy. Medical Image Analysis, vol. 46, pp. 15-25.

- **Polfliet, M.**, Hendriks, M.S., Guyader, J.M., ten Hove, I., Mast, H., Vandemeulebroucke, J., van der Lugt, A., Wolvius, E.B. & Klein, S., 2021. Registration of magnetic resonance and computed tomography images in patients with oral squamous cell carcinoma for three-dimensional virtual planning of mandibular resection and reconstruction. International Journal of Oral & Maxillofacial Surgery.

- Verbanck, S., **Polfliet, M.**, Schuermans, D., Ilsen, B., de Mey, J., Vanderhelst, E. & Vandemeulebroucke, J., 2020. Ventilation heterogeneity in smokers: The role of unequal lung expansion and peripheral lung structure. Journal of Applied Physiology, 129(3), 583-590.

- **Polfliet, M.**, Klein, S., Niessen, W.J. & Vandemeulebroucke, J. , 2021. Laplacian eigenmaps as a dissimilarity measure in multimodal groupwise registration. *under review*

- Farotto, D., Lopez, P., Kvasnytsia, M., **Polfliet, M.**, Vandemeulebroucke, J., Wirix-Speetjens, R. & Vander Sloten, J., 2021. Automated Mitral Annulus Annotation over 4D-CT for Transcatheter Mitral Valve Replacement. *under review*

Publications in international conference proceedings

- **Polfliet, M.**, Vandemeulebroucke, J., Van Gompel, G., Buls, N., Deklerck, R. & Scheerlinck, T., 2015. Estimation of Hip Prosthesis Migration: A Study of Zero Migration. IFMBE Proceedings Series, vol. 45, pp. 126-129.

- **Polfliet, M.**, Klein, S., Niessen, W.J. & Vandemeulebroucke, J., 2017. Laplacian eigenmaps for multimodal groupwise image registration. SPIE: Medical Imaging 2017: Image Processing, vol. 10133.

- Ceranka, J., **Polfliet, M.**, Lecouvet, F., Michoux, N., De Mey, J. & Vandemeulebroucke, J., 2016. Registration Strategies for Whole-Body Diffusion-Weighted MRI Stitching. Computational Diffusion MRI: MICCAI Workshop, Munich, Germany, October 9th, 2015, pp. 195-206.

- **Polfliet, M.**, Klein, S., Huizinga, W., De Mey, J. & Vandemeulebroucke, J., 2016. The Pythagorean averages as group images in efficient groupwise registration. 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). pp. 1261-1264
- Ceranka, J., **Polfliet, M.**, Lecouvet, F., Michoux, N. & Vandemeulebroucke, J., 2017. Whole-body diffusion-weighted MR image stitching and alignment to anatomical MRI. Medical Imaging 2017: Image Processing. vol. 10133.

Conference abstracts

- Scheerlinck, T., **Polfliet, M.**, Deklerck, R., Van Gompel, G., Buls, N. & Vandemeulebroucke, J., 2016. Automated and Marker-Free CT-Based Spatial Analysis Method (CTSA) to Quantify Hip Stem Migration. Development and Validation. International Society for Technology in Arthroplasty (ISTA). pp. 1056-1058.
- Scheerlinck, T., **Polfliet, M.**, Deklerck, R., Van Gompel, G., Buls, N., Van Royen, K. & Vandemeulebroucke, J., 2015. Marker-Free and Automated Computed Tomography Based Spatial Analysis Tool to Quantify Hip Arthoplasty Stem Migration: Development and Validation. Hip International, vol. 25, pp. S37-S37.

Acknowledgements

This PhD project and thesis did not result from an individual effort, but was achieved collectively together with the people closest to me. Where some offered guidance, scientific advice or expertise, others, more importantly, offered emotional support, a helping hand or loyal encouragements.

First of all, I wish to thank my advisors Jef Vandemeulebroucke and Stefan Klein. Jef, this journey started with you. Thank you for offering this academic opportunity to me and for guiding me in it, every single step of the way. Teaching me how to write a proper text or prepare a presentation was a painstaking process, but well worth it. Your dedication to ensure the completion of this project is something I am especially thankful for. Stefan, this journey would not have been possible without you. Your honest feedback and critique were not only insightful and educational, but also motivating and energizing. The guidance you offered on all levels, be it mathematical, algorithmic or scientific, was what made this PhD project possible. For all of this support and guidance, I reiterate my gratitude to both of you. Wiro Niessen, you offered me the opportunity to join EMC and continue this project as a joint effort between our two universities, for which I am grateful. Your guidance helped me in keeping a steady course through this project. I would also like to express my gratitude towards the members of the jury for reviewing this manuscript and providing constructive feedback on its quality.

My thanks also go out to the many people I was able to work with throughout these years on all different kinds of topics. Most importantly, my fellow PhD students and brothers-in-arms, Jakub, Benyameen and Pieter. Many thanks for

the hours and hours of interesting conversation and fun. My gratitude also goes out to the other people of ETRO, Johan, Rudy, Jenny, Panagiotis, Bart, Lubos, Katka, Jennifer. Thanks to all clinical partners in the University hospital, Jette with whom I collaborated, Thierry, Sylvia, Nico, Gert, Toon and to Wyke, Kasper and many other colleagues at Erasmus Medical Center.

My gratitude goes out to all my friends, for all the good times and the distractions offered throughout this project. I wish to thank especially Pieter and Jelle, whom I have always been able to count on and reach out to when I needed it.

A very special thanks to my parents, Rudy and Katrien, who went to great lengths to offer me with a world of possibilities and opportunities. For guiding me and always believing in me, but also for letting me struggle where needed, my sincere gratitude goes out to both of you. I would not be the person I am, without the two of you. Many thanks also to my two brothers, Thomas and Vincent, for shaping me into the person I am today.

Finally, I cannot express with words how grateful I am to have such a loving and supporting partner. Sarah, this accomplishment is as much yours as it is mine, as I am sure your worries and concerns over this project vastly exceeded mine. Your unwavering support and unconditional love throughout this project served as two beacons, safely guiding me to its finish. For this, my deepest gratitude is yours. Now and always.

CHAPTER 10

About the author

Mathias was born on March, 2nd 1989 in Vilvoorde, Belgium. After attending high school at the Sint-Theresia College in Kapelle-Op-Den-Bos, he pursued his interests in physics and astronomy. In 2007, Mathias started the bachelor program in 'Physics and Astronomy' at Ghent University and its ensuing master program. His major graduation work, in the form of a master thesis, was conducted at the department of astronomy of Ghent University in collaboration with the University of Liège and its department of astrophysics. Supported by Michaël Gillon and Maarten Baes, Mathias successfully defended his research on 'Transit timing analysis of the hot Jupiters WASP-43b and WASP-46b and the super Earth GJ1214b'. In 2012, he graduated the master program in 'Physics and Astronomy' with distinction, with biophysics and astronomy as his main interests.

After spending a year teaching physics and mathematics in several high schools, Mathias continued his passion for teaching at the Vrije Universiteit Brussel as a research assistant and enrolled in the PhD program at the faculty of engineering sciences. Here, he would pursue a PhD in medical image analysis, guided by Jef Vandemeulebroucke. In July 2014, he met Stefan Klein at WBIR London, from which a very fruitful collaboration ensued and led to this joint PhD thesis with the Erasmus University in Rotterdam.

On April, 26th 2018, after the birth of his son Victor, Mathias left his research

position in academia and joined ArcelorMittal as a data science consultant, while continuing his work on the PhD thesis. In December 2018, he joined TomTom as a computer vision expert employing multimodal registrations and in August 2020 he joined ROOV, a fintech startup, as a data scientist.

2022 Uitgeverij VUBPRESS
VUBPRESS is an imprint of ASP nv (Academic and Scientific Publishers)
Keizerslaan 34
B-1000 Brussels
Tel. +32 (0)2 289 26 50
Email: info@aspeditions.be
www.aspeditions.be

ISBN 9789461173423
NUR 950
Legal deposit D/2022/11.161/058