Advanced Medical Image Registration Methods for Quantitative Imaging and Multi-Channel Images

Jean-Marie Guyader

The research described in this thesis was carried out at the Erasmus MC – University Medical Centre Rotterdam, the Netherlands.

Financial support for the publication of this thesis was kindly provided by the Erasmus MC and by its Department of Radiology and Nuclear Medicine.

Cover design and all illustrations by Jean-Marie Guyader.

ISBN 978-90-9031-299-6
Printed by Cloître Imprimeurs, Saint-Thonan, Brittany, France.

Advanced Medical Image Registration Methods for Quantitative Imaging and Multi-Channel Images

Geavanceerde medische beeldregistratie methoden voor kwantitatieve beeldvorming en meerkanaals beelden

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. R.C.M.E. Engels

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
dinsdag 22 januari 2019 om 13.30 uur

door

Jean-Marie Guyader
geboren te Guingamp, Frankrijk

Erasmus University Rotterdam

Promotiecommissie

Promotor:	**Prof.dr. W.J. Niessen**
Overige leden:	**Dr. J.A. Hernandez Tamames**
	Prof. J.A. Schnabel
	Prof. A. Jackson
Copromotor:	**Dr.ir. S. Klein**

Contents

General introduction

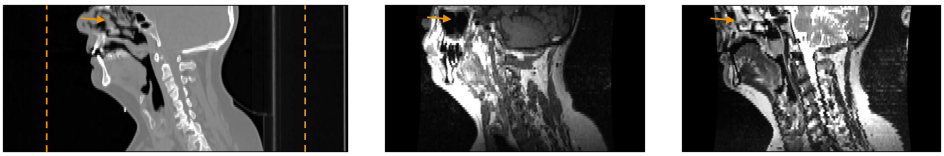
With the development of medical imaging, there has been a growing interest in combining information from multiple images on a voxel-by-voxel basis. This is a challenging task because spatial alignment between different images is often lacking. The process of finding spatial correspondence between two or more images is called *image registration*. In this thesis, we present novel automatic image registration methods.

1.1 Needs for image registration

Spatial correspondence between images is required in many medical applications, but it is rarely fulfilled in practice. There are multiple reasons for that.

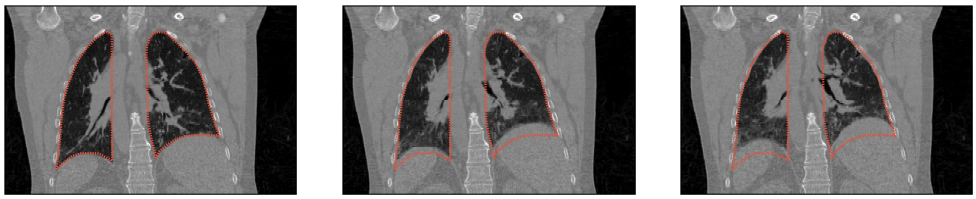
Firstly, images may be acquired using different imaging modalities or imaging devices, possibly inducing a lack of coherence in the coordinate systems that are used to store the images. This is the case in the example shown in Figure 1.1.

Secondly, spatial correspondence between images is not ensured even when they are acquired with the same imaging device. At different acquisition points, the ima-



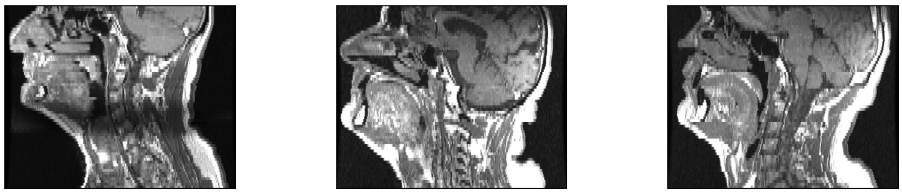
(a) CT (b) T1-weighted MR (c) T2-weighted MR

Figure 1.1: *Images of the head and neck region acquired from the same subject with different imaging modalities. Matrix dimensions and voxel sizes differ between the computed tomography (CT) image (a) and the magnetic resonance (MR) images (b and c). The image space outside the two dashed lines drawn on the CT image is not present in the MR images. Spatial correspondence between the images is not present, as evidenced by the orange arrow pointing at the nasal cavity region.*



(a) Time point 1 (b) Time point 2 (c) Time point 3

Figure 1.2: *Three CT images acquired from the thorax of the same patient at three successive time points. An approximate delineation of the lung, made on the image of the first time point (a), is repeated on the two others (b and c). The positional differences between the images are primarily attributed to respiratory motion in the illustrated case.*



(a) Subject 1 (b) Subject 2 (c) Subject 3

Figure 1.3: *Three T1-weighted MR images obtained from three different subjects. The lack of spatial correspondence between the images is due to the differences in morphology and in posture of the three subjects.*

ged subject may have been positioned in different postures. And even if such bulk motion is taken care of, motion due to breathing or heart beating, for instance, remains a possible source of misalignments (Figure 1.2). Furthermore, health conditions may evolve between successive image acquisitions, which may induce changes in the imaged tissues (e.g. tumour remission, tissue swelling), and therefore impede spatial correspondence between images.

Thirdly, datasets of images may include images acquired from different subjects, which one may wish to register for the purpose of atlas building [26]. In this scenario, spatial correspondence between the acquired images is lacking due to the differences in anatomy and posture of the subjects (Figure 1.3).

The various misalignment sources that we have just cited may be combined in a given dataset.

1.2 Background on image registration

Extensive surveys on medical image registration can be found in the literature [27,88, 117,124,148]. The growing variety of medical imaging datasets has gone along with the parallel development of a considerable number of image registration techniques. Image registration techniques may be classified in various manners.

A first very broad way to characterise an image registration method consists in determining whether it is *feature-based* or *intensity-based*. Feature-based registration approaches are usually based on points, lines or surfaces extracted from the images, and aim at minimising the distance between corresponding features in the images. This requires the extraction of salient features as well as the estimation of their correspondences. The reader is referred to [117] for an overview of feature-based registration methods. The disadvantage of feature-based registration is that any error during the feature extraction stage will propagate into the registration and cannot be recovered at a later stage. Other methods, referred to as intensity-based, use the image intensities directly without the need for feature extraction. These methods measure the degree of shared information between the images, based on the voxel intensities.

A second characterisation criterion consists in determining whether a registration method is *parametric* or *non-parametric*. In parametric image registration, the number of possible transformations is limited by introducing a parametrisation of the transformation used to register the images. For example, a three-dimensional rigid transformation has six parameters (three translation parameters and three rotation parameters). For non-parametric registration techniques, a dense displacement field is estimated which describes the deformation at every voxel. The reader is referred to [39, 96, 117] for an overview of non-parametric methods.

This thesis will focus on registration methods that are both *intensity-based* and *parametric*.

1.3 Image registration components

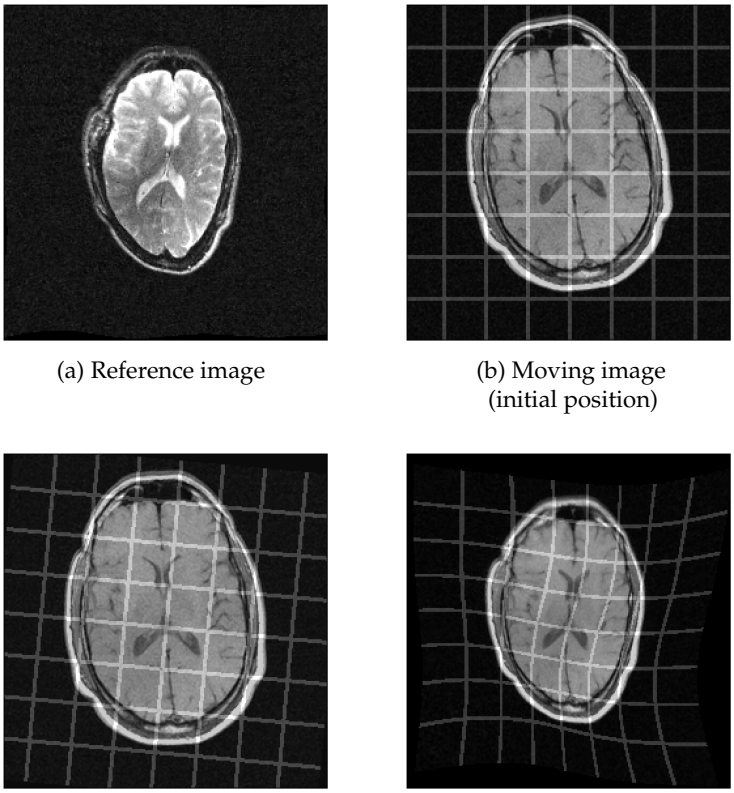
To apply image registration, three main registration components have to be selected.

The first one is the *dissimilarity measure*, denoted $\mathcal{D}$. It quantifies the dissimilarity between the images to register. The choice of dissimilarity measure generally depends on the acquisition modalities of the images to register. There exists a large variety of dissimilarity measures. Examples of the most common dissimilarity measures used for image registration are the sum of squared differences, the correlation coefficient and the mutual information [27]. Dissimilarity measures may be defined for two images, but also in the more general case in which two or more images are considered. For instance, one possible extension of the sum of squared differences for two or more images is the variance dissimilarity measure proposed by Metz et al. [94].

The second registration component is the *transformation model*, which can be rigid or non-rigid. Rigid transformations involve translation and rotation components only. Contrary to rigid registration, non-rigid registration allows to take into account shrinkage or more local deformations. Affine or B-spline transformation models are examples of commonly used non-rigid transformations [117]. Figure 1.4 provides example of images obtained after rigid and non-rigid registration.

The third registration component is the *optimisation technique*, which aims to find the optimal transformation by minimising the measure of dissimilarity between the images. Iterative optimisation procedures are commonly applied to determine that

transformation. Well-known instances of such optimisers are gradient descent [99], quasi-Newton [32], and nonlinear conjugate gradient [28].



(a) Reference image

(b) Moving image (initial position)

(c) Moving image after a rigid registration

(d) Moving image after a non-rigid registration

Figure 1.4: *Example of registration results obtained with two transformation models. The reference image (a) is taken as reference for the registration of the moving image (b). Rigid registration allows rotations and translations, but no shrinking, as shown in (c). Non-rigid registration based on B-splines (d) allows shrinking and deformations. The overlayed grid (not used during registration) provides an indication of the transformations applied to the moving image.*

1.4 Image registration schemes

Multiple strategies have been proposed for the production of sets of registered medical images. These strategies, also referred to as image registration schemes, specify the manner in which the images are registered. For instance, some image registration schemes are based on the selection of a fixed reference image to which the remaining images of the dataset (considered as moving images) are registered, while

other registration schemes do not require the definition of such a fixed reference image. Differences between registration schemes may also involve the number of images that are considered in the optimisation procedures. Commonly, dissimilarity measure optimisation is performed for pairs of $G = 2$ images, but more general methods have been proposed that allow to handle $G \geq 2$ images at a time. This section proposes a categorisation of registration schemes, based on the number of images to register ($G = 2$ images or the general case for $G \geq 2$ images) and describes common registration schemes for these two families of schemes.

1.4.1 Registration schemes for G $=$ 2 images

Most image registrations are formulated as optimisation problems for which the information of two images M_1 and M_2 are considered at a time. This is referred to as *pairwise registration*.

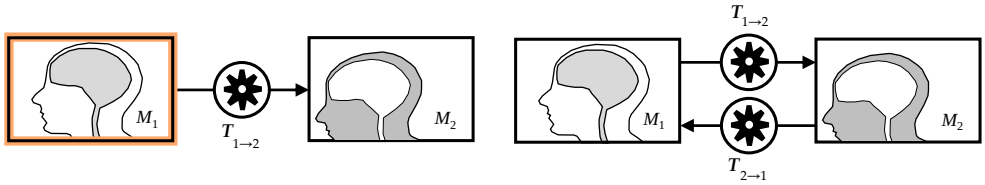
Pairwise registration with a reference image By far, the most common pairwise registration framework consists in choosing a fixed reference image, to which the remaining image (referred to as moving image) is spatially aligned (Figure 1.5a). The aim of pairwise registration is to yield a transformation $\boldsymbol{T}_{1\to 2}$ that maps point coordinates from the image space of the fixed reference image M_1 to the image space of the moving image M_2. One of the disadvantages of pairwise registration with a reference image is that registration results depend on the choice of reference image [45].

Symmetric pairwise registration Other pairwise registration techniques do not require the selection of a reference image: they are both symmetric and pairwise. This is, for instance, the case in the method of Seghers et al. [120], which actually perform two pairwise registrations that alternatively consider the images as fixed reference and moving image (Figure 1.5b). This symmetric pairwise scheme yields transformations $\boldsymbol{T}_{1\to 2}$ and $\boldsymbol{T}_{2\to 1}$, which are combined into transformations that bring the original images M_1 and M_2 into an average image space that is not the image space of any of them. Other examples of symmetric pairwise registration methods include approaches proposed by Avants et al. [9] and Vercauteren et al. [144].

1.4.2 Registration schemes for G $\geq$ 2 images

The two common registration schemes for $G = 2$ images based on pairwise registration described in Section 1.4.1 can be extended to the general case of the registration of $G \geq 2$ images. A third registration scheme, called *groupwise registration*, allows to register $G \geq 2$ images simultaneously in a single optimisation procedure.

Pairwise registration with a reference image The most conventional way for applying image registration to a set of $G \geq 2$ images is to select one fixed reference



(a) Pairwise registration with a fixed image (orange box) and a moving image

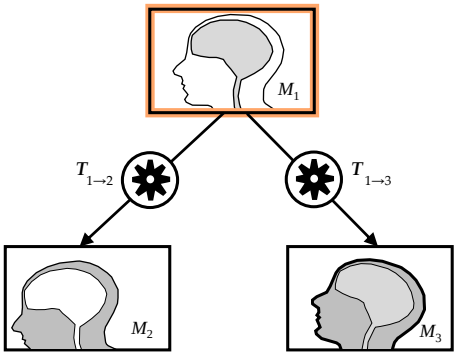
(b) Symmetric pairwise registration

Figure 1.5: *Common registration schemes for $G = 2$ images. Each optimisation procedure is represented by a circled gear symbol.*

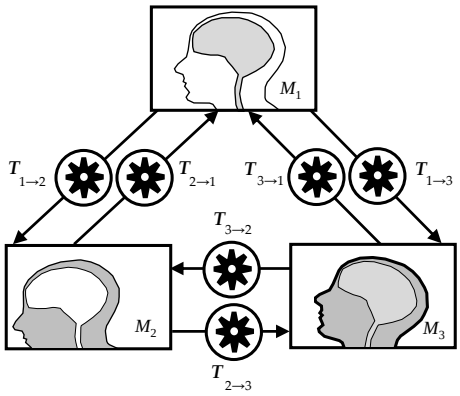
image among them, and then successively register the remaining images to that reference image in an individual manner. This registration scheme is based on the pairwise framework illustrated in Figure 1.5a. Such a pairwise scheme for the registration of a set of $G \geq 2$ images is not always ideal, for two main reasons. The first reason is that pairwise registration requires the selection of a reference image among the images that have to be registered. The selection of the reference image is not always obvious, and may have an impact on the registration results [45]. The second reason is that pairwise registration implies running multiple optimisation procedures in which only two images of the complete image dataset take part. The combination of all image information in a single optimisation procedure is therefore not possible in this pairwise scheme. The pairwise registration scheme with a reference image is illustrated in Figure 1.6a in the case of $G = 3$ images.

Symmetric pairwise registration Similarly to the case with $G = 2$ images, symmetric pairwise registration schemes can be used to register $G \geq 2$ images. The method of Seghers et al. [120] allows to bring the G images to an average image space that is not the image space of one of the original images. This registration scheme does not, however, allow to register all images in a single optimisation procedure. This symmetric pairwise registration scheme is illustrated in Figure 1.6b in the case of $G = 3$ images.

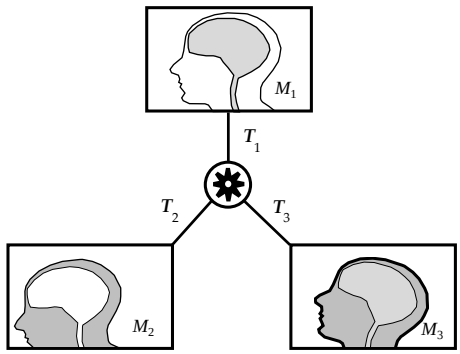
Groupwise registration Groupwise registration schemes are registration schemes that allow the registration of $G \geq 2$ images in a single optimisation procedure, which is not possible with the other registration schemes for $G \geq 2$ images. Groupwise registration schemes allow to take into account all image information in that single optimisation procedure, while the other registration schemes consider only the information of a pair of images during each pairwise registration. Most groupwise registration techniques do not require the selection of a reference image space. A single groupwise registration yields G transformations $\boldsymbol{T}_g$, $g = 1...G$, each of which is applied to the corresponding original image M_g so that is it brought to an image space in which all images are aligned. Such a scheme is illustrated in Figure 1.6c in the case of $G = 3$ images.



(a) Pairwise registration with a reference image.
The orange box indicates that M_1 is taken as reference image.



(b) Symmetric pairwise registration.



(c) Groupwise registration.

Figure 1.6: *Common registration schemes for $G \geq 2$ images.*

1.5 Purpose of this thesis

This thesis proposes advanced medical image registration methods for applications that can be grouped in two broad themes.

The first theme focuses on registration techniques increasing the reliability of *quantitative measurements* extracted from sets of medical images (Figure 1.7). As explained in Section 1.1, multiple factors contribute to the lack of spatial coherence between images, which may have an impact on quantitative measurements extracted in a voxel-per-voxel manner from sets of images. In **Chapters 2**, **3** and **4**, we propose advanced image registration methods that aim at improving the reliability of voxelwise quantitative measurements, based on the registration schemes presented in Section 1.4. The domains of application considered in this theme include quantitative diffusion measurements extracted from diffusion-weighted MR images (DW-MRI), from T1-weighted MR images, and from dynamic contrast-enhanced (DCE) imaging.

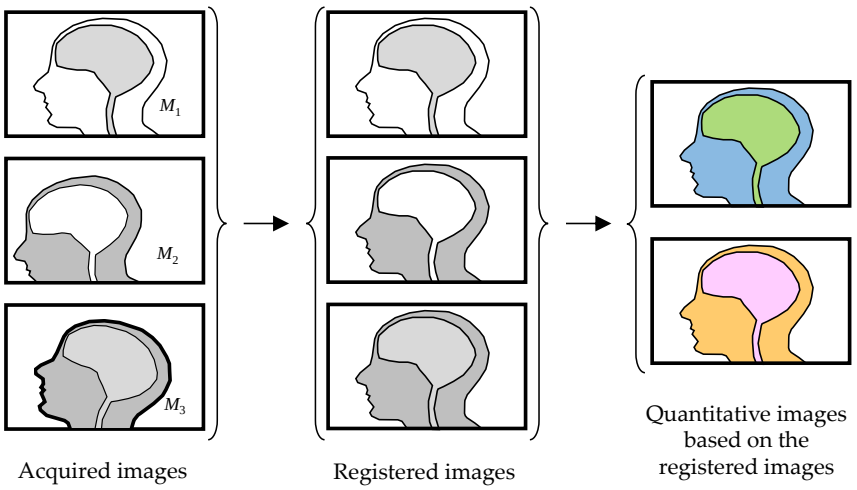


Figure 1.7: *Image registration for quantitative imaging. In the illustrated example, three initial images (denoted M_1, M_2 and M_3) are registered to ensure the reliability of the quantitative images that are extracted from them.*

The second theme that is considered in this thesis is the registration of *multi-channel* images. In medical applications, it happens that the images that have to be registered are composed of multiple channels. The channels of a given multi-channel image may be obtained from different post-acquisition operations (e.g. filtering, computation of feature images) or from different acquisitions (e.g. different modalities or time points). Establishing spatial correspondence between multiple sets of multi-channel images is called multi-channel image registration (Figure 1.8). Pairwise registration techniques have previously been proposed for that problem [113]. In **Chapter 5**, we address the problem of multi-channel image registration by proposing a novel groupwise registration scheme analogous to the one presented in Figure

1.6c but adapted to the multi-channel nature of the images.

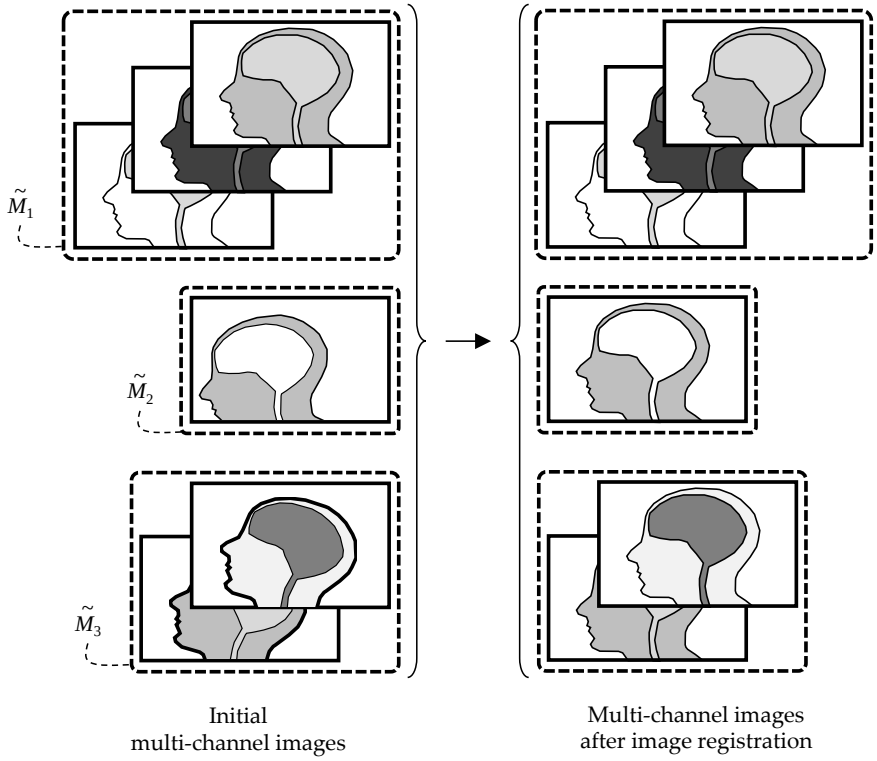


Figure 1.8: *Image registration for multi-channel images. In the illustrated example, the registration of three multi-channel images (denoted $\widetilde{M}_1$, $\widetilde{M}_2$ and $\widetilde{M}_3$) is considered.*

1.6 Chapters overview

This section provides a brief summary of the chapters of this thesis.

Chapter 2 Apparent diffusion coefficient (ADC) images are quantitative parametric maps obtained by applying a curve fitting procedure to multiple DW-MR images. Due to patient motion during their acquisition, it is not ensured that the DW-MRIs are spatially aligned, which may affect the reliability of the ADC images. This chapter develops a pipeline based on automatic three dimensional (3D) non-rigid pairwise and symmetric pairwise image registrations to compensate for misalignments both within each DW-MRI and between all DW-MRIs acquired for a given subject. Evaluation of the method is performed based on ADC images obtained from abdominal free-breathing DW-MRIs.

Chapter 3 This chapter aims to study respiratory tract exacerbation (RTE), based on DW-MR images acquired from patients with cystic fibrosis (CF). Quantitative images are extracted from the DW-MR images based on intravoxel incoherent motion (IVIM), a bi-exponential model yielding quantitative images for molecular-based diffusion, perfusion and for volume fraction. The extraction of quantitative information is preceded by an image registration technique, which like in Chapter 2, ensures spatial correspondence within and between the DW-MRIs. We subsequently assess whether the extracted quantitative IVIM parameters could be used to monitor treatment response during respiratory tract exacerbation in patients with CF.

Chapter 4 In this chapter, we design a dissimilarity measure that can be used for the groupwise registration of $G \geq 2$ images in a single optimisation procedure. Given the widespread use of mutual information for pairwise registration, this chapter proposes to use a multivariate version of mutual information, called total correlation, in the context of groupwise registration. We provide justifications for choosing total correlation as groupwise dissimilarity measure, among other multivariate versions of mutual information. To test the ability of groupwise total correlation to handle multiple numbers of images, the experimental setting involves six types of quantitative MR and dynamic imaging datasets containing between $G = 5$ and $G = 160$ images to register at a time.

Chapter 5 In Chapter 5, we propose a novel groupwise registration framework for the registration of multi-channel datasets of medical images. The key idea is to formulate multi-channel registration as a groupwise image registration problem. The method that we propose is fully modular in terms of dissimilarity measure, transformation model, regularisation method, and optimisation strategy. Besides, it is applicable to any number of multi-channel images, any number of channels per image, and it allows to put in correspondence any pair of images and not just corresponding channels.

Chapter 6 We conclude the thesis with a brief summary, a discussion of general strengths and limitations of our work, and suggestions for future research.

Influence of image registration on apparent diffusion coefficient images computed from free-breathing diffusion MR images of the abdomen

CHAPTER 2

Abstract — Apparent diffusion coefficient (ADC) images are quantitative images that are obtained by applying a curve fitting procedure to multiple diffusion-weighted MR (DW-MR) images. Spatial correspondence between the DW-MR images is not ensured due to patient motion during image acquisition. The curve fitting models used to derive ADC images assume spatial coherence of the DW-MR images. If that condition is not fulfilled, the reliability of the obtained ADC images may be degraded. In this chapter, we evaluated the importance of using image registration techniques to ensure spatial correspondence of the DW-MR images before generating ADC images. We acquired DW-MR images from the abdominal region of free-breathing healthy volunteers. To assess ADC reproducibility, multiple acquisitions of all DW-MR images were performed (two time points, four image series per time point). The image registration pipeline that we developed is based on automatic three-dimensional non-rigid registrations that compensate for motion both within each image and between all images acquired for a given subject. ADC distributions are compared with and without image registration in abdominal volumes of interest. Besides, the effects of interpolation and Gaussian blurring as alternative trivial strategies to reduce motion artefacts are also investigated. Among the four considered scenarios (no processing, interpolation, blurring and registration), registration yields the best alignment scores. In particular, ADCs obtained without registration are 30% higher than with registration, based on the considered datasets. Registration improves voxelwise reproducibility at least by a factor of 2 and decreases uncertainty (Fréchet-Cramér-Rao lower bound). Registration provides similar improvements in reproducibility and uncertainty as acquiring four times more data.

Based upon: J.-M. Guyader, L. Bernardin, N. H. M. Douglas, D. H. J. Poot, W. J. Niessen and S. Klein, "Influence of image registration of ADC images computed from free-breathing diffusion MRIs of the abdomen", *Journal of Magnetic Resonance Imaging*, vol. 42, no. 2, pp. 315–330, 2015.

2.1 Introduction

The apparent diffusion coefficient (ADC) is a non-invasive measure providing quantitative information on the diffusion of water molecules in biological tissues [74]. Pathophysiological processes such as cancer are known to have an impact on cell density, which translates into diverse diffusion properties. This is the reason why the ADC constitutes a potentially interesting imaging biomarker in the field of oncology drug development. ADC images can be computed from diffusion-weighted MR images (DW-MRIs) characterised by different b-values and diffusion gradient directions. This chapter focuses on abdominal ADCs, with a particular interest in the liver.

Spatial alignment of the acquired DW-MRIs is not guaranteed if the subject moves during the acquisition [58, 100]. Misalignments may be due to patient bulk motion. In the abdominal region, misalignments are particularly prone to occur because of respiratory and cardiac motion, inducing poor image quality. The issue of image quality is commonly addressed by acquiring each DW-MRI several times during an imaging session and averaging them [57, 66, 68, 101]. Despite improving the signal-to-noise ratio (SNR) of the resulting ADC image, this technique does not compensate for motion. It also causes blurring and leads to longer acquisition times. A first alternative to averaging consists of preventing motion during the acquisition by means of breath holding, triggering or gating [17, 57, 61, 70, 71, 131, 132]. These methods have the advantage of addressing the issue of motion at the source, therefore reducing the need for image post-processing. However, breath holding requires a short scan time. Also, triggering and gating do not always perform well if the respiratory rhythm is irregular [19]. A second alternative to averaging is post-acquisition motion compensation [6, 66, 82, 90, 102, 103]. In this chapter, we use image registration as a post-acquisition motion compensation technique. Its goal is to bring all acquired DW-MRIs into a common image space using 3D deformable transformations and to subsequently extract ADC quantitative images. According to the acquisition protocol, individual DW-MRIs may be affected by misalignments. In the general case, motion should therefore be compensated at the level of individual images (intra-image registration), but also motion between the various images should be compensated (inter-image registration).

We therefore propose an image registration pipeline that brings all DW-MRIs of a given patient into a common image space, using both intra-image and inter-image registration. The method is quantitatively evaluated on ten abdominal imaging datasets of five healthy volunteers using a free-breathing protocol (two scanning sessions per subject). The ADC images obtained after applying our motion compensation pipeline to the DW-MRIs are compared to ADC images obtained without applying motion compensation. Results obtained for two alternative scenarios to image registration are also provided. Evaluation is based on the computation of uncertainty and reproducibility measures.

For ADC computations from diffusion MRIs of the abdomen and thorax, the closest related works involving post-acquisition motion compensation based on image registration are the studies of Arlinghaus et al. [6] and Mazaheri et al. [90]. These articles include an image registration step and compare ADC maps without and with

registration. In this chapter, we follow a similar approach for abdominal images. Compared to existing work, the main contributions of this chapter are the following: first, data analysis is performed using not only regionwise but also voxelwise measurements, allowing to take into account organ heterogeneity. Second, this chapter includes measures of the precision of the estimated ADCs based on the square root of the Fréchet-Cramér-Rao lower bound, denoted FCRLB_σ. Third, we propose to study reproducibility of the ADCs at two levels. Since each volunteer is scanned on two occasions, the baseline and follow-up scans can be used for assessing inter-visit reproducibility. The acquisition protocol also provides the opportunity to study intra-visit reproducibility. In order to ensure consistency of the volumes of interest, segmentations are propagated to the various series and scanning sessions using image registration. Finally, we investigate to what extent the duration of the acquisition could be reduced if image registration was used.

2.2 Materials and methods

2.2.1 Subjects

Five healthy volunteers (volunteer 1 – sex: female, age: 30 years old, body mass index: 19.5 kg/m²; volunteer 2 – male, 35 years, 27.7 kg/m²; volunteer 3 – female, 64, 27.1 kg/m²; volunteer 4 – male, 63, 23.1 kg/m²; volunteer 5 – male, 62, 29.3 kg/m²) were scanned twice in a fasted state, with the same imaging set-up and protocol. The average time between the two scanning sessions was 7 days (time range: 3 to 12 days). The study was conducted with the approval of the ethics committees of the participating institutes.

2.2.2 Acquisition protocol

Diffusion-weighted MRIs were acquired on a 1.5 T MR scanner (MAGNETOM Avanto; Siemens Healthcare, Erlangen, Germany), using a multi-slice 2D echo-planar imaging (EPI) sequence in the transverse orientation. Repetition time (TR) was 8,000 ms and echo time (TE) 95 ms. Matrix size was $256\times224\times40$ and 40 slices were acquired with an in-plane spatial resolution of 1.48×1.48 mm², slice thickness (d) 5 mm, field of view (FOV) 38×38 cm^2, bandwidth 1,776 Hz/px and EPI factor 112. Neither respiratory nor cardiac triggering were used and a SPAIR (spectral attenuated inversion recovery) fat suppression was applied. The duration of each scanning session was 16.7 minutes. The subjects were not asked to hold their breath. A free-breathing protocol was selected based on the advantages reported in literature [66], among which are: more flexible sequence design, greater choice of b-values, better patient compliance, and possibility of performing multiple slice excitations. In addition, ADCs in the liver have been proven not to be significantly different according to the selected type of protocol: free-breathing or breath-holding [70].

For each volunteer and each scanning session, four successive series of images were acquired. For each series, 28 image type were acquired, an image type being defined in this chapter by a condition of b-value and diffusion gradient direction.

The number of images acquired per volunteer and per scanning session was therefore 112. A graph showing the acquisition timeline is provided in Figure 2.1. Out of the 16.7 minutes of the scanning session, only 14.9 are dedicated to the actual image acquisition. The images are denoted $I_{b,\boldsymbol{g},s}$, with indexes b, $\boldsymbol{g}$ and s respectively referring to the b-value, gradient direction and series index (s = 1...4). Ten b-values b were used: 0, 50, 100, 150, 200, 300, 500, 900, 1200 and 1600 s/mm^2. The diffusion gradients $\boldsymbol{g}$ were successively set along three orthogonal directions **x**, **y** and **z**, with **x** and **y** defining the axial plane, except for b = 0 s/mm^2 (no specific orientation). Each three-dimensional image $I_{b,\boldsymbol{g},s}$ was reconstructed from 40 two-dimensional slices, with the particularity that they were not acquired contiguously: the odd slices were acquired first in the inferior-superior direction, followed by the even slices, in the same direction. The consequence is that two consecutive slices in the reconstructed volume were acquired 4 seconds apart. This interleaved acquisition scheme is meant to reduce cross-talk between slices [13].

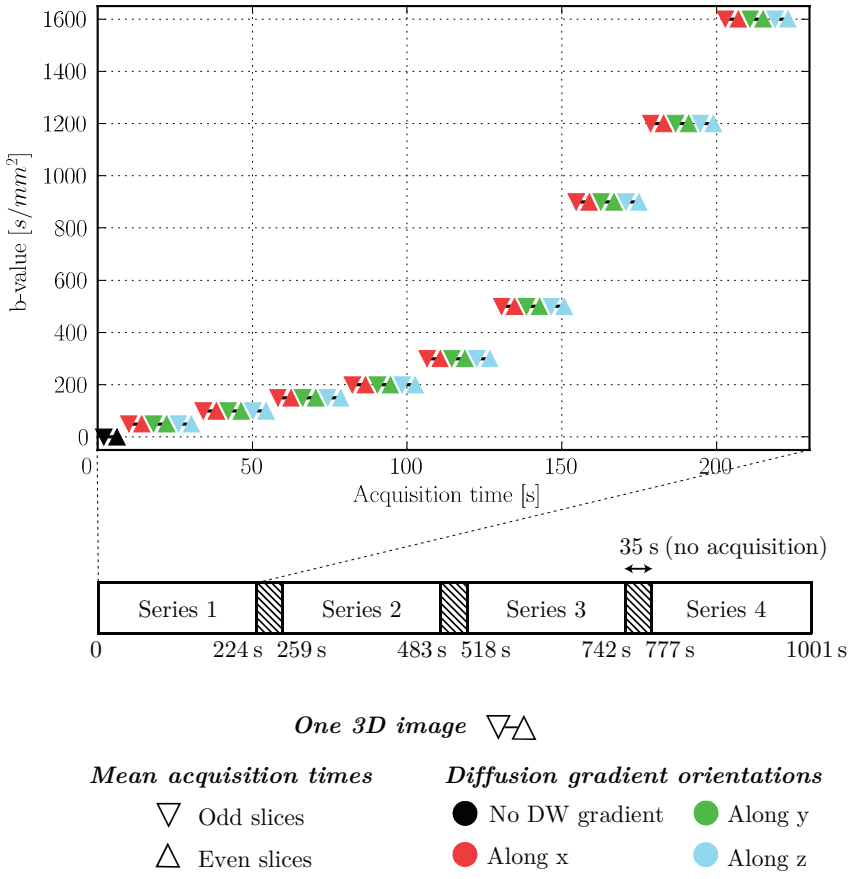


Figure 2.1: *Acquisition sequence for one scanning session. A ▽ symbol represents the mean acquisition time of the odd slices (△ for the even slices).*

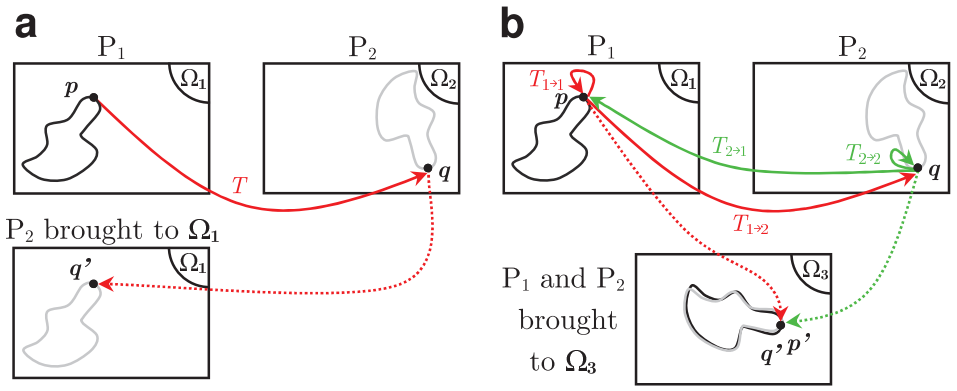


Figure 2.2: *Pairwise registration scheme with a reference image (a) and symmetric pairwise registration scheme (b). In the pairwise case with reference image, a moving image (P_2) is aligned to a fixed reference (P_1). In the symmetric pairwise case, all images are considered as moving images: each image to all other images and the computed transformations are combined to bring the images to a common mid-point space in which they are aligned.*

2.2.3 Background: pairwise registration with a reference image and symmetric pairwise registration

Image registration is commonly applied using a pairwise scheme with a reference image (Figure 2.2a). Two images are considered [56]: a fixed reference image P_1 and a moving image P_2, with their respective image spaces Ω_1 and Ω_2, and $\boldsymbol{x}$ the spatial coordinate. This scheme consists in finding the transformation $T : \Omega_1 \rightarrow \Omega_2$ that spatially aligns $\mathrm{P}_2(T(\boldsymbol{x}))$ with $\mathrm{P}_1(\boldsymbol{x})$. Using T, P_2 is then brought to Ω_1. This pairwise scheme is not well suited to alignment problems for which there is no obvious reference image. In such cases, other registration schemes can be employed.

In this section, we propose to use a symmetric pairwise registrations scheme based on multiple pairwise registrations [120]. Let us consider n images P_i, $i = 1...n$ to be aligned with such a method. Figure 2.2b provides an example for $n = 2$ images. For each i, P_i is taken as fixed image and n independent registration are performed between each P_j, $j = 1...n$, and P_i, yielding n transformations $T_{i \rightarrow j}$. Each P_i is then resampled into an average or mid-point image space using $\bar{T}_i^{-1}(\boldsymbol{x})$, the inverse of the arithmetic mean of the transformations $T_{i \rightarrow j}$, with $i = 1...n$.

$$\bar{T}_i^{-1}(\boldsymbol{x}) = \left(\frac{1}{n} \sum_{j=1}^{n} T_{i \rightarrow j}(\boldsymbol{x}) \right)^{-1} \tag{2.1}$$

Both pairwise with reference and symmetric pairwise schemes are used in this chapter. Unlike the pairwise schemes, the groupwise schemes require only one optimisation procedure to register two or more images. Groupwise methods like the method of Metz et al. [94] were not investigated in the framework of this chapter.

2.2.4 Description of the image processing pipeline

The automatic pipeline that we designed consists of several registration steps, followed by a ADC curve fitting step. All registrations are performed using a 3D non-rigid transformation model and a mutual information dissimilarity measure [133]. The various steps are described in this section and illustrated in Figure 2.3.

2.2.4.1 Step ① – Intra-image registration

Due to respiratory motion, the interleaved acquisition scheme described in Section 2.2.2 creates artefacts between the odd and even slices of the acquired images $\mathrm{I}_{b,g,s}$ (Figure 2.4a). The first step of our image processing pipeline is therefore to compensate for such odd/even artefacts. From each acquired DW-MRI $\mathrm{I}_{b,\boldsymbol{g},s}$, two three-dimensional subvolumes $\mathrm{I}_{b,\boldsymbol{g},s,\mathrm{odd}}$ and $\mathrm{I}_{b,\boldsymbol{g},s,\mathrm{even}}$ are extracted, respectively based on the odd slices and on the even slices, centered at their original positions but with a doubled slice thickness. $\mathrm{I}_{b,\boldsymbol{g},s,\mathrm{odd}}$ and $\mathrm{I}_{b,\boldsymbol{g},s,\mathrm{even}}$ therefore do not have empty lines between two of their immediately neighbouring slices. The hypothesis is made that individual odd or even subvolumes are not affected by motion artefacts. This intra-image registration step consists of applying the symmetric pairwise registration technique described in Section 2.2.3 to the two subvolumes $\mathrm{I}_{b,\boldsymbol{g},s,\mathrm{odd}}$ and $\mathrm{I}_{b,\boldsymbol{g},s,\mathrm{even}}$. Once the two subvolumes are registered, they are resampled in 3D to the resolution of the original image $\mathrm{I}_{b,\boldsymbol{g},s}$. This resampling process includes interpolations in the 3D space. The voxelwise average of the resampled motion-corrected odd and even subvolumes is finally computed, yielding the images $\mathrm{J}_{b,\boldsymbol{g},s}$.

2.2.4.2 Steps ② to ④ – Inter-image registration

The images $\mathrm{J}_{b,g,s}$ (step ①) are brought into a common image space in steps ② to ④.

Step ② – ***Symmetric pairwise registration of the four repeated scans*** As mentioned in Section 2.2.2, each type of image is acquired four times during a scanning session. The goal of step ② is to register the images of each set of four repeated scans. The four intra-corrected images $\mathrm{J}_{b,\boldsymbol{g},s}$, with $s = 1...4$, are registered in a symmetric pairwise manner (Section 2.2.3) for each pair $(b, \boldsymbol{g}) \in B \times G$. This yields the transformations $\{R_{b,\boldsymbol{g},s}, s = 1...4\}$ and the registered images $\{\mathrm{K}_{b,\boldsymbol{g},s}, s = 1...4\}$, which are then averaged in a voxelwise manner to yield an image $\overline{\mathrm{K}}_{b,\boldsymbol{g}}$ with an improved SNR.

Step ③ – ***Pairwise registration between b-values images*** Spatial correspondence between the various image types is established in this step. The voxelwise average $\overline{\mathrm{K}}_{0,\mathbf{0}}$ is chosen as fixed image because it has the highest SNR. 27 independent pairwise registrations are performed with the $\overline{\mathrm{K}}_{b\neq 0,\boldsymbol{g}\neq \mathbf{0}}$ as moving images, producing the transformations $T_{b,\boldsymbol{g}}$. The mutual information dissimilarity measure that we used is particularly suitable for aligning images characterised by different intensity distributions, which is the case when DW-MR images of different b-values are considered. In addition, the non-rigid transformation model that was used during registration is rather conservative, which avoids that low SNR images be deformed in a too extreme manner.

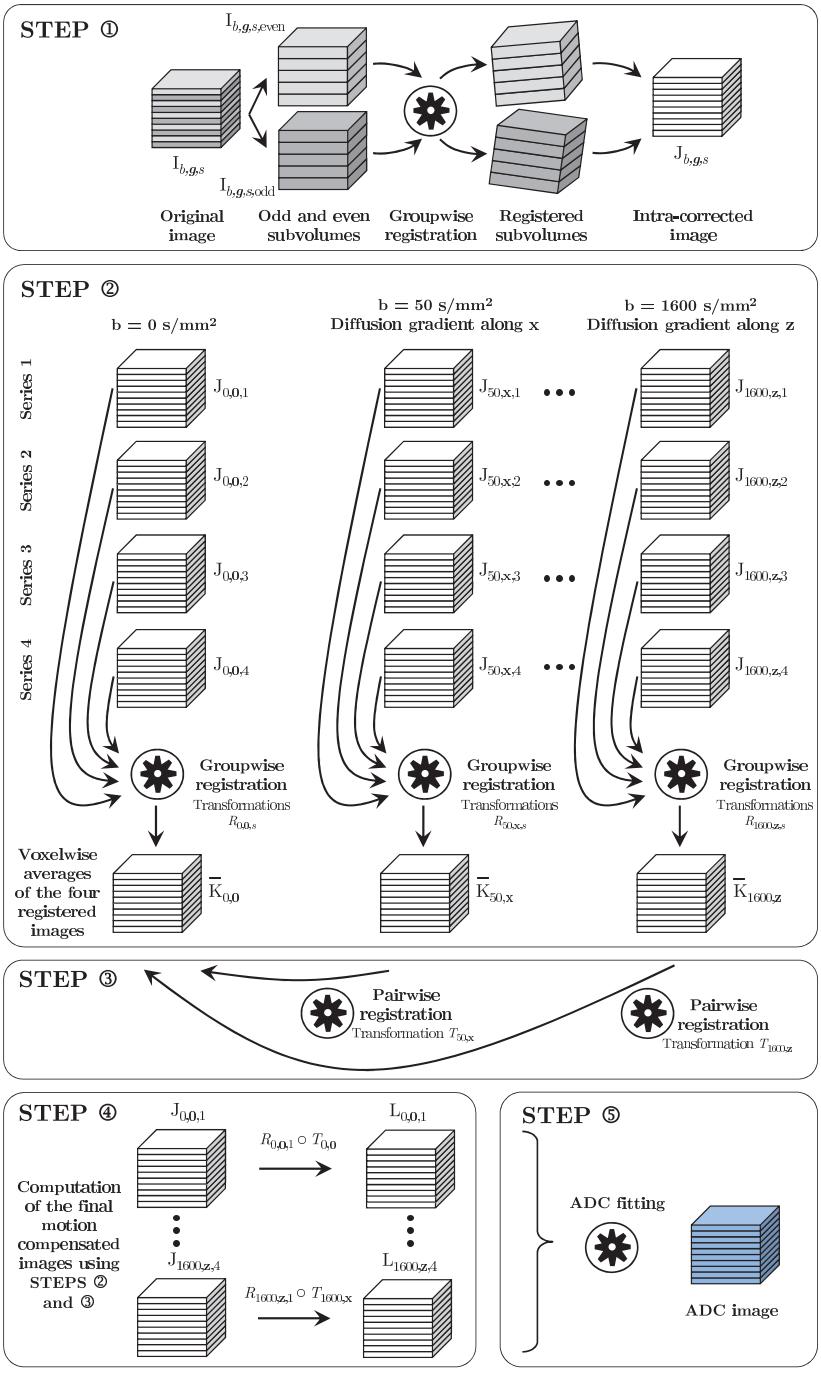


Figure 2.3: *Image processing pipeline. Step ①: intra-image registration. Steps ②, ③ and ④: inter-image registration. Step ⑤: maximum likelihood ADC fitting.*

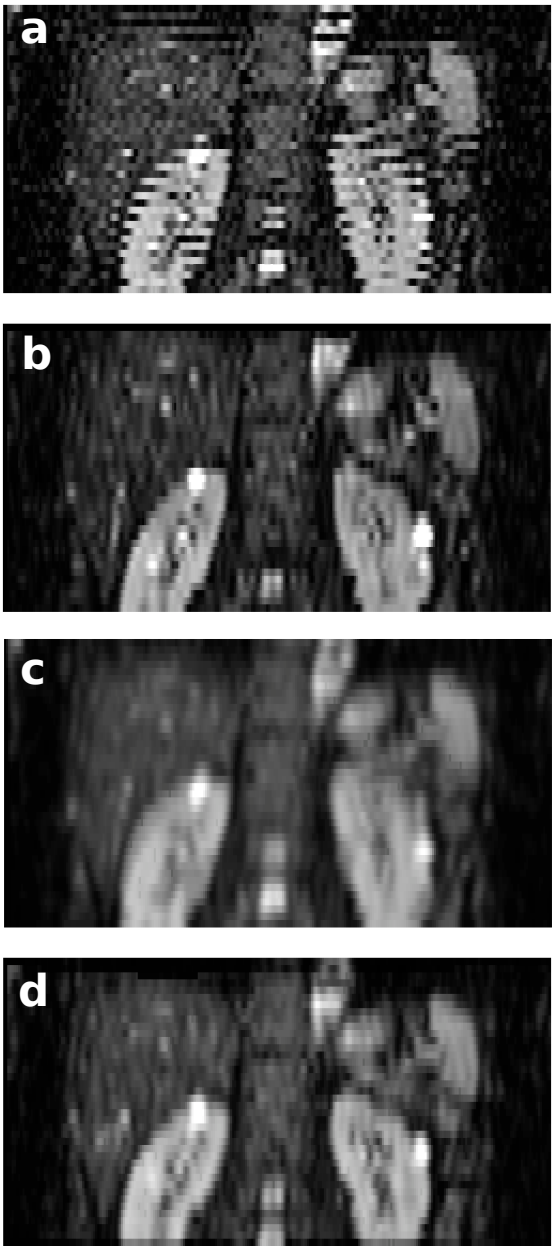


Figure 2.4: *Coronal b = 0 s/mm^2 DW-MR images. 'No processing' scenario (a), 'interpolation' scenario (b), 'Gaussian blurring' scenario (c), and 'image registration' scenario (d).*

Step ④ – *Final resampling* A simple final step to bring all the images into the same image space would be to successively apply the transformation $R_{b,\boldsymbol{g},s}$ to the intra-aligned images $\mathrm{J}_{b,\boldsymbol{g},s}$ and then the $T_{b,\boldsymbol{g}}$ to the $\mathrm{K}_{b,\boldsymbol{g},s}$ images. As this would imply interpolating the images twice, we choose instead to apply the composite transformation $R_{b,\boldsymbol{g},s} \circ T_{b,\boldsymbol{g},s}$ to the intra-compensated images $\mathrm{J}_{b,\boldsymbol{g},s}$ to obtain the final images $\mathrm{L}_{b,\boldsymbol{g},s}$. The ADCs are then extracted by curve fitting from the $\mathrm{L}_{b,\boldsymbol{g},s}$ (Section 2.2.4.3). The $\overline{\mathrm{K}}_{b,\boldsymbol{g}}$ images are only used for estimating $T_{b,\boldsymbol{g}}$.

2.2.4.3 Step ⑤ – ADC computation

Knowledge of the complete diffusion tensor D is not necessary for computing ADCs, which is why only three distinct diffusion gradient directions were used in the acquisition. A mono-exponential model is used to extract the diagonal elements of the diffusion tensor D, alongside with an estimation of s_0, the MR intensity without diffusion gradient. s is the observed MR intensity for a given b-value b (in s/mm^2). Each gradient direction is described by its unit vector $\boldsymbol{g}$ (Equation 2.2):

$$s = s_0 \exp\left[-b\left(\boldsymbol{g}^{\mathrm{T}} D \boldsymbol{g}\right)\right] \tag{2.2}$$

The ADC is defined as the average of diagonal elements of the diffusion tensor D (Equation 2.3):

$$\mathrm{ADC} = \frac{d_{1,1} + d_{2,2} + d_{3,3}}{3} \tag{2.3}$$

An MR-specific curve fitting technique based on Poot et al. [107] is used to extract ADC values from the acquired DW-MRIs. It consists of a maximum likelihood estimator that takes into account the Rician noise characteristics in magnitude MR data [122]. The optimisation problem thus contained the four parameters: s_0, $d_{1,1}$, $d_{2,2}$ and $d_{3,3}$. The Fréchet-Cramér-Rao lower bound (FCRLB) is in addition computed at each spatial location and provides a lower bound of the variance of the ADC. The square root of FCRLB, denoted FCRLB$_\sigma$, will be reported because it has the same dimension as the ADC. The FCRLB$_\sigma$ indicates the theoretical uncertainty of the ADC value computed at a given spatial location.

Previous studies showed that micro-circulation and perfusion effects cause deviation from the mono-exponential model for low b-values (i.e. under 50 – 100 s/mm^2) [5,33,67,71–73]. In order to avoid the influence of perfusion, images acquired with b = 0 and 50 s/mm^2 were not used in the fitting procedure. The DW-MRIs with b = 1200 and 1600 s/mm^2 were also not used due to their too low signal-to-noise ratios. The curve fitting was consequently performed on the images $\mathrm{L}_{b,\boldsymbol{g},s}$, with $b \in$ {100, 150, 200, 300, 500, 900} s/mm^2, $\boldsymbol{g} \in G$ and s = 1...4.

2.2.4.4 Reproducibility assessment

The fact that four repeated series of measurements (s = 1...4) were acquired for each visit allows to study the intra-visit reproducibility of the computed ADCs. This is made possible because the scanner parameters remained unchanged during one visit. It is therefore of interest to apply image registration to each series taken independently and compare the ADC results over the four series. To process a single

series ('1 series' case), a simpler version of the pipeline is designed, in which the inter-image registration step only consists in registering the $J_{b\neq 0,\boldsymbol{g}\neq \mathbf{0},s=s_0}$ DW-MRIs of the considered series s_0 with $J_{b=0,\boldsymbol{g}=\mathbf{0},s=s_0}$ (this replaces steps ②, ③ and ④).

2.2.5 Software

The post-acquisition motion correction pipeline that is described in this chapter is based on elastix [65], a publicly available open source image registration software. Optimisation is performed using stochastic gradient descent [64], a multi-resolution strategy with 2 resolution levels and a maximum number of 2,000 iterations per resolution. Mutual information [87, 105, 133] is chosen as dissimilarity measure as it is particularly suitable for handling registration across images with different intensity distributions, such as DW-MR images acquired with different b-values and/or diffusion gradient directions. A three-dimensional B-spline transformation model [118] with control points spacing of 64 mm is utilised to describe the motion of the patients during the acquisition. Such a conservative point spacing is meant to avoid that images with a low SNR be deformed in a too extreme fashion. The parameter files used for the registrations are available on the webpage of elastix[1]. Image manipulations, including format conversion, sorting, preparation of the registrations as well as the management of the registration steps is performed using Python scripts (version 2.7.3) alongside with the following additional packages: NumPy 1.6.2, SciPy 0.11.0, pydicom 0.9.7 and NiBabel 1.3.0. Some functions of the open source Insight Toolkit [56] were also used for converting image formats. The ADC fitting was carried out using MATLAB.

2.3 Experiments

2.3.1 Considered scenarios

The DW-MRIs and ADC images obtained with the pipeline described in Section 2.2 are compared with three other scenarios. In total, four scenarios were considered in this chapter.

The first, referred to as 'no processing', consists in applying the ADC curve fitting directly on the acquired images $I_{b,\boldsymbol{g},s}$. In a second scenario called 'interpolation', the even slices of the original DW-MR images are extracted to form new volumes. Linear interpolation between the even slices is then performed to deduce the odd slices. This scenario simulates the images that would have been obtained if the interleaved acquisition protocol had not been chosen. The third scenario, denoted 'registration', considers the images $L_{b,\boldsymbol{g},s}$ obtained after applying the proposed image registration pipeline. In addition to compensating for motion, image registration also introduces blurring owing to interpolation. It is however not clear what the effect of blurring is. A fourth scenario is therefore introduced, referred to as 'Gaussian blurring'. It consists of applying a three-dimensional Gaussian kernel to the acqui-

[1] http://elastix.bigr.nl/wiki/index.php/Parameter_file_database

red images $I_{b,g,s}$. A standard deviation of 1 voxel for the Gaussian kernel was used, as it was sufficient to make the odd/even artefacts visually disappear (Figure 2.4b).

2.3.2 Experiment 1 – Motion compensation accuracy

The goal of this first experiment is to quantify the alignment accuracy in the four scenarios described in Section 2.3.1. For that purpose, the whole spleen is manually delineated on several DW-MRIs: for each volunteer, the images characterised by $b \in \{0, 100, 500, 900\}$ s/mm², $G = \{\boldsymbol{x}\}$ and $s \in \{1, 4\}$ are segmented. For the 'interpolation' scenario, the segmentations are performed after interpolation of the odd slices on the original image space, for the 'Gaussian blurring' scenario, on the DW-MRIs after blurring, and for the 'registration' scenario, after applying the intra and inter-registration steps. Dice similarity coefficients are subsequently computed with respect to the $b = 0$ s/mm^2 images. The spleen was chosen because it is an organ located under the lungs in an area subject to respiratory motion and because it is relatively easy to segment. Paired t-tests are used to compare the Dice coefficient distributions obtained in the four scenarios.

In the 'registration' scenario, the processed DW-MR images are also visually examined to check whether they show unreasonable deformations, such as pronounced stretchings and twistings.

2.3.3 Experiment 2 – Quantitative ADC analysis

2.3.3.1 Volumes of interest

The second experiment is dedicated to the comparison of ADC values obtained with the four scenarios. For each volunteer, two 3D spherical volumes of interest (VOI) were defined with a radius of 15 mm on the first non diffusion-weighted image $I_{0,\mathbf{0},1}$ in the 'no processing' scenario. The same VOI was used in the 'interpolation', 'Gaussian blurring' and 'image registration' scenarios. Such a sphere encompasses approximately 1000 voxels. The first of the two VOIs (Figure 2.5a) is positioned in a homogeneous region of the right lobe of the liver. Given the fact that the effects of image registration are expected to be more visible in non-homogeneous regions, a second VOI was selected in a nearly sub-hepatic area encompassing both the right lobe of the liver and the top of the right kidney (Figure 2.5b).

The ADC images obtained with the four scenarios are first compared visually, and subsequently quantitatively analysed in terms of median value, homogeneity (using the interquartile range of the ADC distribution), uncertainty (using the $FCRLB_\sigma$), and in terms of reproducibility across series and visits. Paired t-tests are used to compare the distributions of the median ADC values obtained with the four scenarios.

VOIs delineated on the scans of the first visit $I_{0,\mathbf{0},1}$ are not usable for the second visit because the coordinate system of these images are generally different. In the 'registration' scenario, the VOIs were propagated using registration to allow for a consistent comparison between series and scanning session. The chosen solution

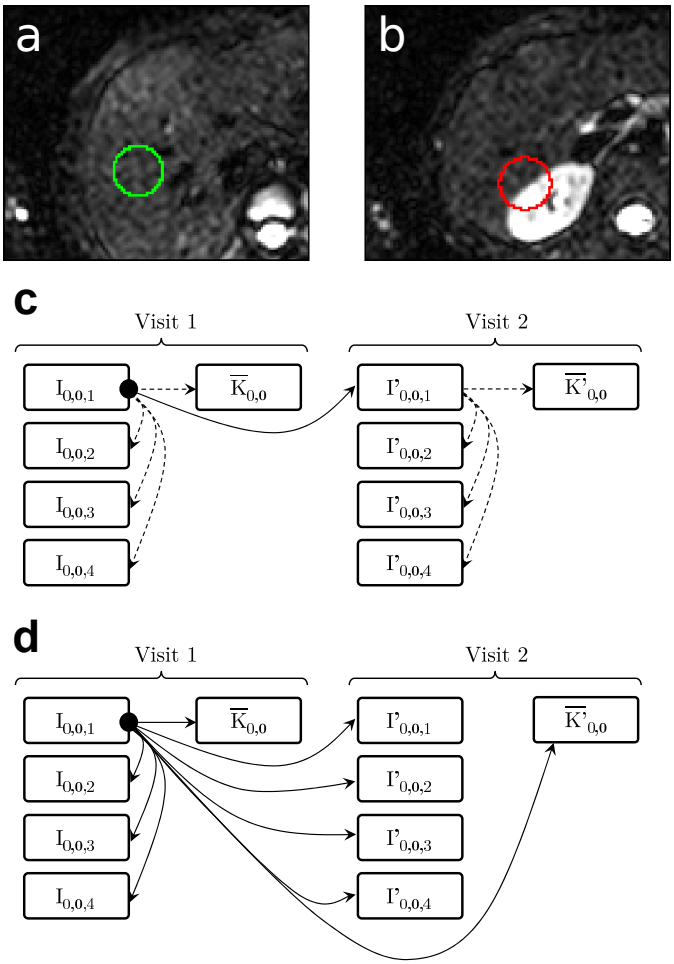


Figure 2.5: *Example of VOIs overlayed on a* b *= 0 s/mm^2 DW-MRI. The first VOI is placed in the right lobe of the liver (a), and the second between the liver and the right kidney (b). In the 'no processing', 'interpolation' and 'Gaussian blurring' scenarios (c), the VOIs are directly propagated (dashed arrow: $\dashrightarrow$), while they are propagated using registration (full arrow: $\longrightarrow$) in the 'registration scenario' (d).*

consists of registering the first non diffusion-weighted image of the second visit ($I'_{0,\mathbf{0},1}$) to the first non diffusion-weighted image of the first visit ($I_{0,\mathbf{0},1}$). This allows to propagate the VOI from the first visit to the second. For the 'interpolation', 'no processing' and 'Gaussian blurring' scenarios, the VOI considered for the first series of each visit is also used for the other series (Figure 2.5c). In the 'image registration' scenario (Figure 2.5d), all the non diffusion-weighted MRIs from both visits are brought to the image space of ($I_{0,\mathbf{0},1}$), which is followed by a propagation of the VOI in each case. The propagation of VOIs allows comparisons of the ADC distribu-

tion within each scanning session (intra-visit reproducibility) and between the two scanning sessions (inter-visit reproducibility or baseline/follow-up reproducibility).

2.3.3.2 Voxelwise and regionwise analyses

Two complementary approaches are used to quantitatively compare ADCs.

Voxelwise approach The first approach, called voxelwise, consists of calculating a standard deviation value at each spatial location, from multiple images. The first objective was to quantify intra-visit reproducibility. To that end, we computed a standard deviation image from the four ADC images (one for each series s) obtained for a given scanning session. The second objective was to quantify inter-visit reproducibility. This was realised by computing a standard deviation image from the ADC images obtained for the two scanning sessions. From the standard deviation images, 90$^{\text{th}}$ percentiles are extracted within each VOI, yielding the observed voxelwise variability measures 'STD intra' and 'STD inter'. In addition, we computed an estimated variability measure based on 90$^{\text{th}}$ percentiles of the FCRLB$_\sigma$ obtained by fitting. Such a measure was previously used in other studies to evaluate the effect of motion on quantitative parameters [16].

Regionwise approach The second approach is regionwise: median ADC values are first computed for each VOI. Standard deviations are then computed, yielding observed regionwise variability measures. In the regionwise case , the observed variability measures may be compared to an estimate of the standard deviation of the median ADC. Given the fact that it is not possible to compute such an estimate analytically, a Monte Carlo experiment was carried out: the N voxels of a VOI were considered as random variables that are independent but non-identically distributed. For each voxel, a new value was generated using the normal distribution $\mathcal{N}(\mu, \sigma)$, with μ the ADC and σ the FCRLB$_\sigma$ at this voxel location. A new median value was then stored. This operation was repeated 10,000 times: the estimate of the standard deviation of the median ADC was obtained by computing the standard deviation of the 10,000 medians.

2.3.4 Experiment 3 – Data averaging

As mentioned in Section 2.2, four analogous series of DW-MRIs are acquired during a scanning session. Multiple series of images are often acquired in the context of DW-MR imaging for improving image quality by averaging [57, 66, 68, 101]. Despite improving the signal-to-noise ratio (SNR) of the resulting ADC images, this technique does not fundamentally compensate for motion.

This experiment focuses on comparing the ADCs obtained by applying curve fitting to the four series of misaligned images taken together ('No processing – All series'), and comparing the obtained results with one series of misaligned images ('No processing – 1 series') and with one series of images aligned with our registration technique ('Registration – 1 series'). The quantities that are compared are: the median ADC, the 90$^{\text{th}}$ percentile voxelwise FCRLB$_\sigma$, 90$^{\text{th}}$ voxelwise STD inter and the IQR.

2.4 Results

2.4.1 Results for Experiment 1 – Motion compensation accuracy

Figure 2.4 compares analogous b = 0 s/mm² DW-MRIs reconstructed in the coronal plane for each of the four scenarios (volunteer 4, visit 1). The inter-slice staircase artefacts are visually removed in the 'interpolation', 'Gaussian blurring' and the 'image registration' scenarios. Figure 2.4 additionally shows that the 'interpolation' and 'image registration' scenarios lead to sharper images, compared to the 'Gaussian blurring' scenario.

In Figure 2.6, the alignments of b = 0 s/mm², b = 100 s/mm², b = 500 s/mm², and b = 900 s/mm² can be compared in the 'no processing' and 'image registration' scenarios, for one of the datasets. In addition to showing that the images are better aligned with registration than without, we also observe that no unrealistic motion is introduced by registration between different b-value DW-MR images, indicating that the chosen mutual information dissimilarity measure is adequate.

In terms of Dice coefficients, Figure 2.7 shows that the overlaps are higher for the 'image registration' scenario than for the 'interpolation', 'Gaussian blurring' and 'no processing' scenarios, with respective mean Dice coefficients of 0.88, 0.86, 0.85, 0.84 for b = 100 s/mm², 0.87, 0.82, 0.85, 0.83 for b = 500 s/mm² and 0.85, 0.81, 0.81, 0.77 for b = 900 s/mm². The results indicate that the Dice coefficients are significantly different only between the 'no processing' and 'image registration' scenarios.

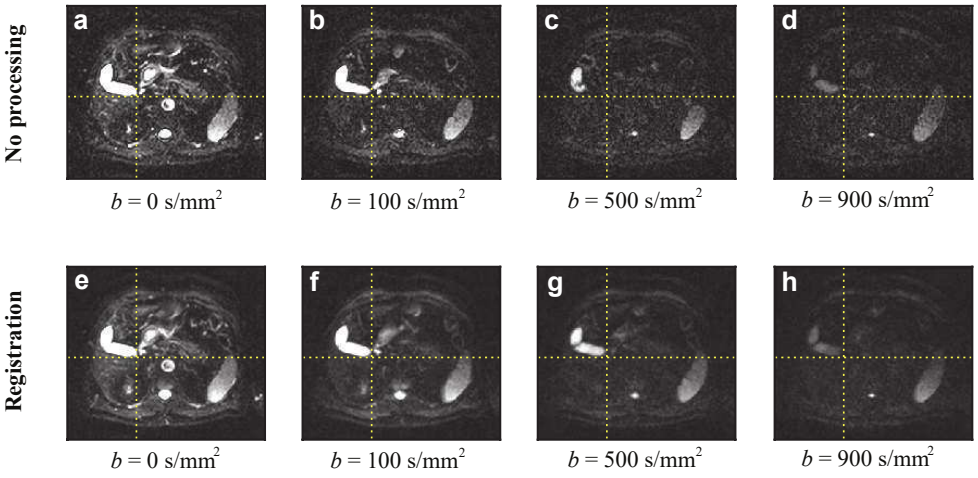


Figure 2.6: *Examples of different b-value DW-MR images in the 'no processing' and 'image registration' scenarios (volunteer 5, first visit). The b = 100 s/mm², b = 500 s/mm² and b = 900 s/mm² images (diffusion gradient direction x) are more similar to the b = 0 s/mm² image in the 'image registration' scenario. The mutual information metric used for registration handles the differences in intensity distribution between the images.*

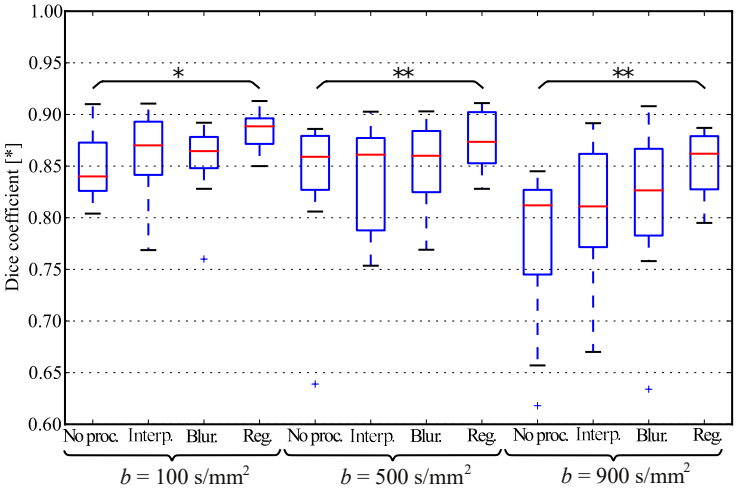


Figure 2.7: *Dice overlap coefficients between manual spleen segmentation obtained on the b = 100, 500, 900 s/mm^2 images and the non diffusion-weighted image (b = 0 s/mm^2). Best overlaps are obtained for the 'registration' scenario. Statistical significance of the paired t-tests between the scenarios is denoted with *: $p < 0.05$, **: $p < 0.001$.*

2.4.2 Results for Experiment 2 – Quantitative ADC analysis

Figure 2.8 provides examples of computed ADC maps for one of the visits (volunteer 1, visit 1), considering two cases: a first case, in which only one series of DW-MRIs was used in the ADC curve fitting, and a second case in which the data of all series was used in the ADC curve fitting. Visual inspection of the ADC images suggests that the 'Gaussian blurring' and 'image registration' scenarios both improve the visual quality of the ADC maps with respect to the 'no processing' scenario: the organs are better visualised and the number of voxels for which the fitting fails decreases. The ADC images computed using Gaussian blurring and image registration visually appear to be rather similar, but a closer inspection shows differences in sharpness of the images (compare Figure 2.8d and 2.8f): the organs look sharper in the 'image registration' scenario than in the 'Gaussian blurring' scenario.

Median values computed from the two VOIs (first VOI: liver, second VOI: interface between the liver and the right kidney) are reported in Tables 2.1 and 2.2. For a given volunteer and a given VOI, median ADCs are quite similar across visits and series. Condensed results corresponding to the average of the median ADCs over the ten visits are presented in Tables 2.1 and 2.2, and shown graphically as boxplots in Figure 2.9. In the 'all series' case, the median ADCs computed on the first VOI are respectively 0.87 $\mu m^2/ms$ without image processing, 0.83 $\mu m^2/ms$ in the 'interpolation' scenario, 0.79 $\mu m^2/ms$ with Gaussian blurring and 0.79 $\mu m^2/ms$ when using the image registration pipeline. For the second VOI, the respective median ADCs are: 1.45 $\mu m^2/ms$, 1.46 $\mu m^2/ms$, 1.43 $\mu m^2/ms$ and 1.22 $\mu m^2/ms$. For both VOIs, the median ADCs obtained in the 'no processing' scenario are always higher compared to

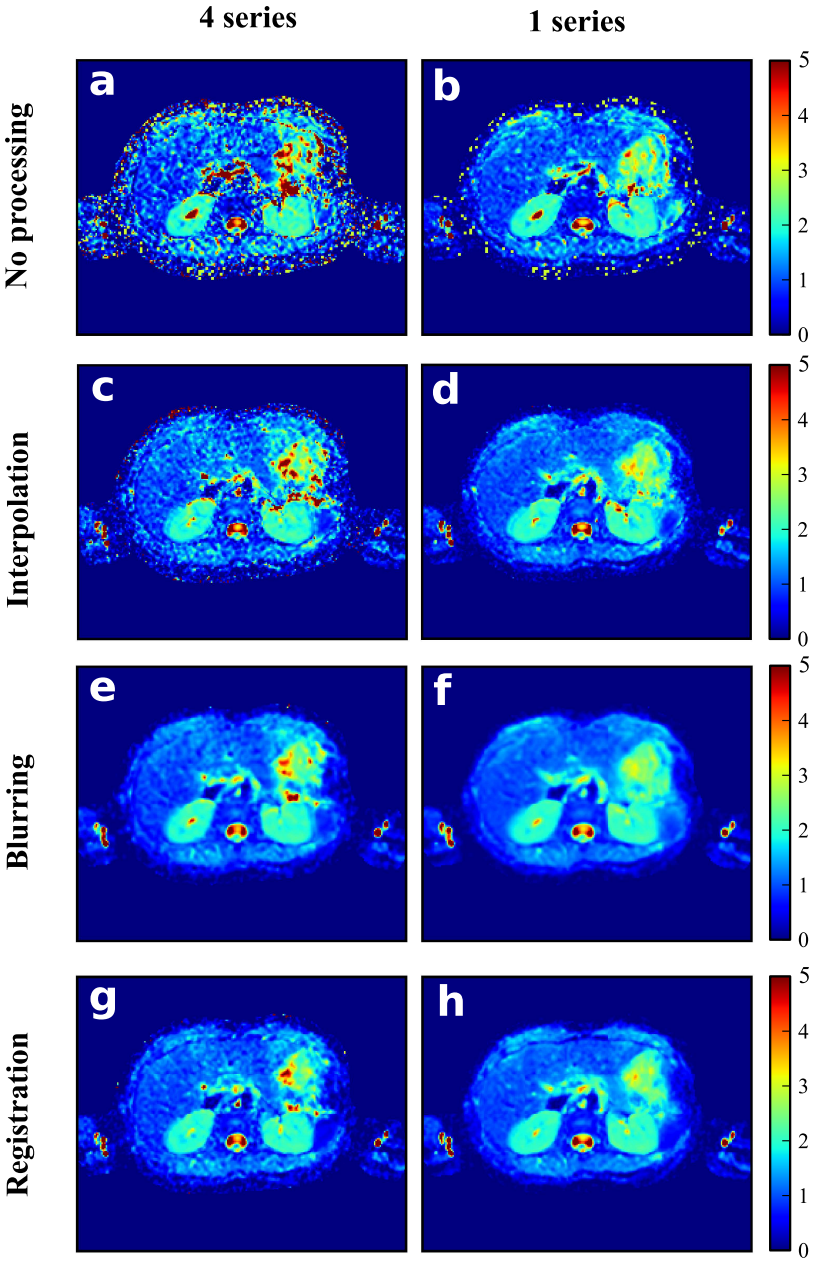


Figure 2.8: *Computed ADC maps. Colour scales are given in* $\mu m^2/ms$.

the 'Gaussian blurring' and 'image registration' scenarios. In the 'all series' case, median ADCs are 9.2% (first VOI) and 19.4% (second VOI) higher in the 'no processing' scenario, compared to the 'image registration' scenario. In the 'Series 1' case, the overestimation is respectively 21.2% and 32.9%. ADC values are also overestimated in the 'interpolation' scenario, but to a lower extent than for the 'no processing' scenario. In the 'Gaussian blurring' and 'image registration' scenarios, median ADCs ('Average' line of Tables 2.1 and 2.2 and Figure 2.9) are more comparable between separate series and all series than in the two other scenarios. For each scenario, the median ADC obtained when using all series is compared with the average of the four median ADCs obtained when only one series is fitted. The absolute difference between these two values divided by the corresponding median ADC in the 'all series' case is 12.8% without registration, 8.1% in the 'interpolation' scenario, 1.4% with blurring and 1.7% with registration for the first VOI (respectively 19.8%, 14.4%, 3.3% and 7.6% for the second VOI). Interquartile ranges (IQRs), characterizing the homogeneity of the ADCs within a given VOI, are lower with blurring or registration than when no processing is applied to the images (Figure 2.9c and 2.9d).

Regionwise and voxelwise reproducibility results are provided in Figure 2.10. In the voxelwise approach (Figures 2.10a and 2.10b), FCRLB_σ, STD intra and STD inter values computed in the 'Gaussian blurring' and 'image registration' scenarios are at least twice as low as for the 'no processing' and 'interpolation' scenarios for both VOIs. The FCRLB_σ are of the same order of magnitude as STD intra and STD inter, but always lower. In the regionwise analysis (Figures 2.10c and 2.10d), STD inter, STD intra and FCRLB_σ are in general reduced in the 'image registration' scenario, with respect to the 'no processing' scenario. Monte Carlo estimates of the median ADC are found to be much lower than the observed variabilities STD inter and intra.

2.4.3 Results for Experiment 3 – Data averaging

Table 2.3 compares the 'no processing – all series' scenario with the 'image registration – 1 series' and 'no processing – 1 series' scenarios. For a given scenario, the results of the table are averaged over all patients. For scenarios focusing on one series, the values correspond to an average of the four individual series.

For both VOIs, the median ADCs, interquartile ranges, 90^{th} percentile FCRLB_σ and 90^{th} percentile STD inter computed in the 'image registration – 1 series' scenario are quite similar to these obtained in the 'no processing – all series'. This is less the case when comparing 'no processing - all series' and 'no processing – 1 series'.

This experiment indicates that considering only one series of registered images yields ADC image characteristics that are quite similar to four series of non-registered images. Besides, these results also suggest that ADCs obtained from non-registered images are overestimated with respect to ADCs obtained from registered images.

Table 2.1: *First VOI: median ADCs. All values are given in $\mu m^2/ms$. The mean and standard deviations (last column) are calculated using the four median values of series 1, 2, 3 and 4. The 'Average' lines contain values averaged over the 10 visits.*

			Series 1	2	3	4	All series	Mean ± std
Volunteer 1	Visit 1	*No proc.*	1.03	1.00	1.20	1.05	1.02	1.07 ± 0.09
		Interp.	0.99	0.96	1.18	1.10	1.03	1.06 ± 0.10
		Blur.	0.98	0.96	1.15	1.00	1.01	1.02 ± 0.09
		Reg.	0.99	0.96	1.13	1.03	1.03	1.03 ± 0.07
	Visit 2	*No proc.*	1.19	1.15	1.13	1.15	1.13	1.16 ± 0.03
		Interp.	1.13	1.11	1.70	1.15	1.13	1.14 ± 0.03
		Blur.	1.16	1.13	1.12	1.13	1.13	1.14 ± 0.02
		Reg.	1.17	1.15	1.14	1.19	1.16	1.16 ± 0.02
Volunteer 2	Visit 1	*No proc.*	1.00	0.95	0.97	1.17	0.92	1.02 ± 0.10
		Interp.	0.89	0.83	0.91	1.14	0.87	0.94 ± 0.14
		Blur.	0.78	0.80	0.81	1.00	0.85	0.85 ± 0.10
		Reg.	0.76	0.79	0.80	1.01	0.81	0.84 ± 0.11
	Visit 2	*No proc.*	0.91	1.06	1.10	1.04	0.94	1.03 ± 0.08
		Interp.	0.85	0.98	0.97	0.98	0.89	0.95 ± 0.06
		Blur.	0.75	0.86	0.91	0.87	0.84	0.85 ± 0.07
		Reg.	0.76	0.88	0.94	0.86	0.83	0.86 ± 0.07
Volunteer 3	Visit 1	*No proc.*	0.69	0.62	0.71	0.69	0.53	0.68 ± 0.04
		Interp.	0.61	0.49	0.37	0.81	0.48	0.57 ± 0.19
		Blur.	0.43	0.47	0.48	0.52	0.47	0.48 ± 0.04
		Reg.	0.46	0.54	0.56	0.49	0.49	0.51 ± 0.05
	Visit 2	*No proc.*	0.72	0.69	0.68	0.64	0.55	0.68 ± 0.03
		Interp.	0.66	0.62	0.55	0.59	0.51	0.61 ± 0.05
		Blur.	0.54	0.54	0.54	0.54	0.53	0.54 ± 0.00
		Reg.	0.55	0.51	0.60	0.53	0.53	0.55 ± 0.04
Volunteer 4	Visit 1	*No proc.*	1.08	1.27	0.99	1.37	1.08	1.18 ± 0.17
		Interp.	1.05	1.26	0.82	1.23	1.05	1.12 ± 0.16
		Blur.	0.94	1.06	0.85	1.12	0.97	0.99 ± 0.12
		Reg.	0.93	1.05	0.85	1.16	0.99	1.00 ± 0.14
	Visit 2	*No proc.*	1.01	1.01	1.14	1.14	1.00	1.08 ± 0.08
		Interp.	0.89	0.95	1.01	1.14	0.94	1.00 ± 0.11
		Blur.	0.83	0.88	0.97	0.95	0.90	0.91 ± 0.06
		Reg.	0.84	0.89	0.98	0.91	0.92	0.91 ± 0.06
Volunteer 5	Visit 1	*No proc.*	1.02	0.87	0.97	0.82	0.72	0.92 ± 0.09
		Interp.	0.81	0.73	0.76	0.68	0.65	0.75 ± 0.05
		Blur.	0.62	0.56	0.60	0.55	0.56	0.58 ± 0.03
		Reg.	0.59	0.60	0.62	0.54	0.56	0.59 ± 0.03
	Visit 2	*No proc.*	0.85	0.97	1.04	0.95	0.76	0.95 ± 0.08
		Interp.	0.76	0.72	0.82	0.94	0.71	0.95 ± 0.09
		Blur.	0.57	0.59	0.67	0.68	0.61	0.63 ± 0.06
		Reg.	0.55	0.62	0.64	0.65	0.60	0.62 ± 0.05
Average		*No proc.*	0.95	0.96	0.99	1.00	0.87	0.98 ± 0.08
		Interp.	0.86	0.87	0.87	0.98	0.83	0.89 ± 0.10
		Blur.	0.76	0.79	0.81	0.84	0.79	0.80 ± 0.06
		Reg.	0.76	0.80	0.83	0.84	0.79	0.81 ± 0.06

Table 2.2: *Second VOI: median ADCs. All values are given in $\mu m^2/ms$. The mean and standard deviations (last column) are calculated using the four median values of series 1, 2, 3 and 4. The 'Average' lines contain values averaged over the 10 visits.*

			Series				All series	Mean ± std
			1	2	3	4		
Volunteer 1	Visit 1	*No proc.*	1.41	1.24	1.62	1.43	1.24	1.43 ± 0.16
		Interp.	1.45	1.25	1.62	1.50	1.30	1.46 ± 0.15
		Blur.	1.38	1.22	1.57	1.44	1.39	1.40 ± 0.15
		Reg.	1.28	1.33	1.32	1.27	1.40	1.30 ± 0.03
	Visit 2	*No proc.*	1.35	1.25	1.37	1.33	1.22	1.33 ± 0.05
		Interp.	0.41	1.26	1.43	1.38	1.28	1.37 ± 0.08
		Blur.	1.28	1.20	1.25	1.26	1.21	1.25 ± 0.03
		Reg.	1.27	1.20	1.28	1.29	1.23	1.26 ± 0.04
Volunteer 2	Visit 1	*No proc.*	2.10	1.58	1.74	1.33	1.43	1.69 ± 0.32
		Interp.	2.13	1.61	1.56	1.55	1.42	1.71 ± 0.28
		Blur.	2.04	1.41	1.53	1.29	1.46	1.57 ± 0.33
		Reg.	1.62	1.40	1.33	1.31	1.25	1.42 ± 0.14
	Visit 2	*No proc.*	1.30	1.53	1.54	1.93	1.26	1.58 ± 0.26
		Interp.	1.33	1.56	1.72	1.81	1.41	1.61 ± 0.21
		Blur.	1.16	1.30	1.56	1.81	1.38	1.46 ± 0.29
		Reg.	1.18	1.30	1.59	1.50	1.44	1.39 ± 0.19
Volunteer 3	Visit 1	*No proc.*	2.21	1.91	1.98	2.02	1.49	2.03 ± 0.13
		Interp.	2.02	1.77	1.79	1.81	1.53	1.85 ± 0.12
		Blur.	1.51	1.52	1.53	1.68	1.48	1.56 ± 0.08
		Reg.	1.37	1.34	1.44	1.45	1.23	1.40 ± 0.05
	Visit 2	*No proc.*	1.87	1.92	1.41	1.41	1.11	1.65 ± 0.28
		Interp.	1.43	1.72	1.31	1.29	1.08	1.44 ± 0.20
		Blur.	1.24	1.33	1.03	1.03	1.03	1.16 ± 0.15
		Reg.	1.09	1.37	1.16	1.12	1.05	1.19 ± 0.13
Volunteer 4	Visit 1	*No proc.*	2.25	2.06	2.18	1.89	1.90	2.10 ± 0.16
		Interp.	1.93	1.81	2.31	1.88	1.85	1.98 ± 0.22
		Blur.	1.89	1.99	1.85	1.69	1.77	1.86 ± 0.13
		Reg.	1.77	1.57	1.65	1.25	1.64	1.56 ± 0.22
	Visit 2	*No proc.*	1.92	1.70	2.22	2.12	1.88	1.99 ± 0.23
		Interp.	1.82	0.56	2.07	2.12	1.87	1.89 ± 0.26
		Blur.	1.82	1.61	1.95	1.91	1.86	1.82 ± 0.15
		Reg.	1.70	1.25	1.58	1.55	1.32	1.52 ± 0.19
Volunteer 5	Visit 1	*No proc.*	1.88	1.55	1.88	1.81	1.29	1.78 ± 0.16
		Interp.	2.16	1.23	1.57	1.66	1.24	1.66 ± 0.38
		Blur.	1.16	1.07	1.43	1.08	1.18	1.19 ± 0.17
		Reg.	1.08	1.10	0.79	0.92	0.85	0.97 ± 0.15
	Visit 2	*No proc.*	1.94	1.79	1.67	2.02	1.72	1.86 ± 0.16
		Interp.	1.91	1.40	1.49	1.93	1.57	1.68 ± 0.28
		Blur.	1.45	1.40	1.32	1.73	1.50	1.48 ± 0.18
		Reg.	1.34	1.12	0.96	0.98	0.77	1.10 ± 0.18
Average		*No proc.*	1.82	1.65	1.76	1.73	1.45	1.74 ± 0.19
		Interp.	1.76	1.52	1.69	1.69	1.46	1.66 ± 0.22
		Blur.	1.49	1.41	1.50	1.49	1.43	1.47 ± 0.17
		Reg.	1.37	1.30	1.31	1.26	1.22	1.31 ± 0.13

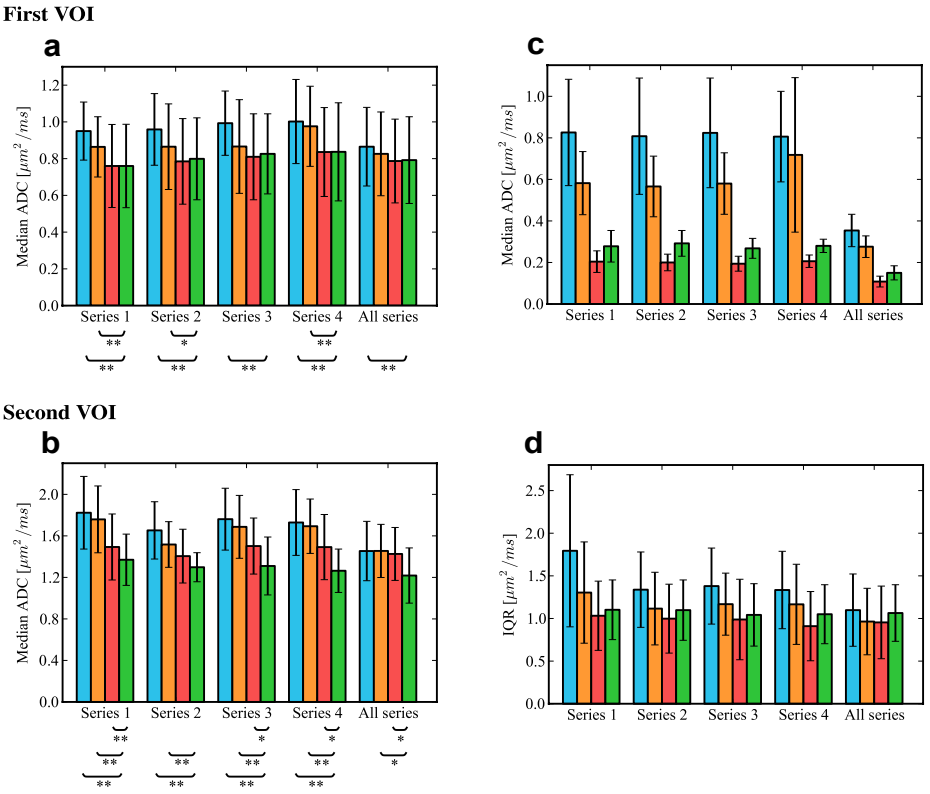


Figure 2.9: *Median ADCs (a,b) and interquartile ranges (IQR) (c,d) averaged over the 10 visits. Blue ■: 'no processing' scenario, orange ■: 'interpolation' scenario, red ■: 'Gaussian blurring' scenario, and green ■: 'image registration' scenario. The error bars represent the standard deviation of the measurements over the visits. Paired t-tests compare the median ADC sets obtained with respect to the image registration scenario (*: $p < 0.05$, **: $p < 0.01$).*

Table 2.3: *Comparison of the median ADCs, interquartile ranges, FCRLB$_\sigma$ and inter-visit voxelwise reproducibilities: no processing, 4 series versus registration, 1 series. Values are given in $\mu m^2/ms$.*

VOI	Measurement	No processing All series	Registration 1 series	No processing 1 series
1	Median ADC	0.87	0.80	0.98
	90$^{\text{th}}$ perc. FCRLB$_\sigma$	0.13	0.11	0.33
	90$^{\text{th}}$ perc. STD inter	0.27	0.24	0.84
	IQR	0.36	0.28	0.80
2	Median ADC	1.45	1.31	1.74
	90$^{\text{th}}$ perc. FCRLB$_\sigma$	0.12	0.11	0.27
	90$^{\text{th}}$ perc. STD inter	0.48	0.39	2.27
	IQR	1.10	1.07	1.46

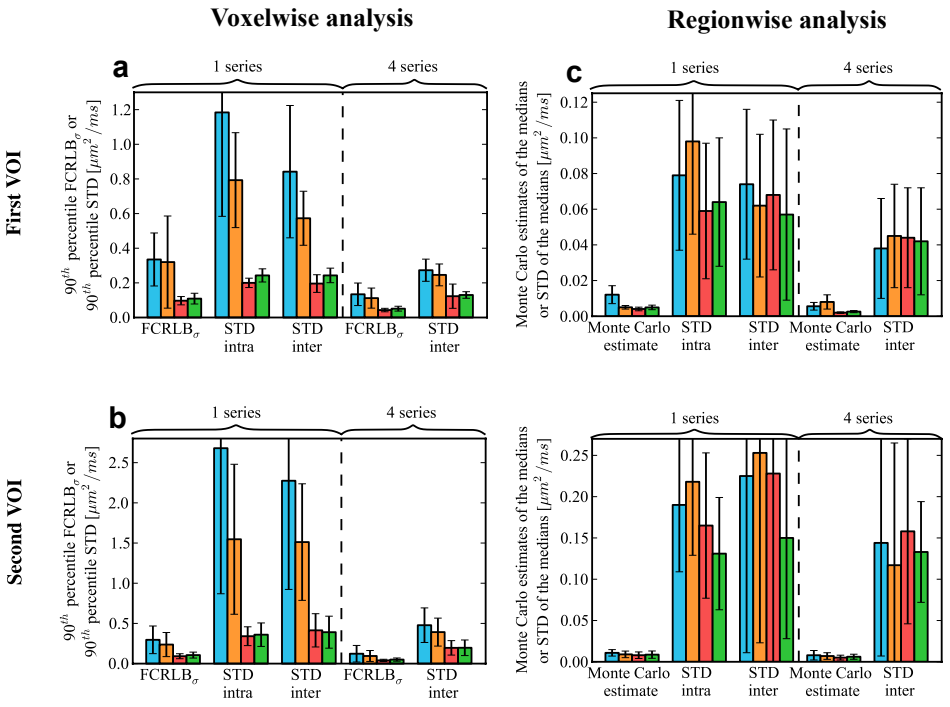


Figure 2.10: *Voxelwise (a,b) and regionwise (c,d) reproducibility analysis. The results are averaged over the 10 visits and the error bars represent the standard deviation of the measurements over the visits. Blue ■: 'no processing' scenario, orange ■: 'interpolation' scenario, red ■: 'Gaussian blurring' scenario, and green ■: 'image registration' scenario.*

2.5 Discussion

In this chapter, a comprehensive image processing pipeline based on image registration is proposed and evaluated for ADC quantification. The main objective is to compare ADC measurements obtained in four scenarios: with our registration pipeline, without this pipeline, with Gaussian blurring of the acquired DW-MRIs, or by replacing the odd slices of the original image volumes by interpolations between the even slices. A full non-rigid approach was investigated in the 'image registration' scenario. In Section 2.2, we made the hypothesis that the odd and even subvolumes $\mathrm{I}_{b,\boldsymbol{g},s,\mathrm{odd}}$ and $\mathrm{I}_{b,\boldsymbol{g},s,\mathrm{even}}$ extracted from the acquired volume $\mathrm{I}_{b,\boldsymbol{g},s}$ are not affected by motion artefacts. This hypothesis is in practice not always true because 4 seconds are required to acquire each 3D subvolume, which is also of the order of magnitude of the respiratory period. Through plane misalignments can thus persist within the subvolumes. This effect is however partly corrected in the intra-registration step thanks to the use of a non-rigid transformation model, though it cannot be claimed that the exact original shapes are recovered.

With respect to the 'no processing' scenario, motion compensation accuracy

reached higher Dice coefficients in the 'interpolation', 'Gaussian blurring' and 'image registration' scenarios but was only significantly higher in the 'image registration' case. This first indicates that the DW-MR images are better aligned in the 'image registration' scenario. This also suggests that the blurring effect due to the interpolations carried out during the resampling steps in the 'image registration' scenario is not sufficient to explain the increase in Dice similarity coefficient: motion compensation plays a clear positive role in realigning the images.

The 'no processing' scenario leads to an overestimation of the ADC values of approximately 30% with respect to the 'image registration' scenario, for the considered datasets. Applying the 'interpolation' scenario also results in an overestimation with respect to the 'no processing' scenario, which indicates that the interleaved acquisition order is not the only cause of overestimation. The 'Gaussian blurring' scenario induces either underestimations (VOI 1) or overestimations (VOI 2) of median ADCs, with respect to the 'image registration' scenario. These elements indicate that the 'image registration' scenario has an influence on ADC quantification that can be explained only by motion compensation. The presented results indicate that 'Gaussian blurring' and 'image registration' both improve visual quality of ADC maps. These two scenarios result in a comparable fit uncertainty ($FCRLB_\sigma$) and reproducibility (STD inter and STD intra). Median ADCs in a large homogeneous liver volume of interest (first VOI) are very similar for these two scenarios. However, this is not the case for an inhomogeneous region located at the interface between the liver and the kidney (second VOI). This indicates that blurred ADC images are biased at the interfaces between organs. In patients with tumors, preserving the sharpness of the images may be essential: mixing the tumor signal with the healthy liver parenchyma may result in quantitative discrepancies.

Both the voxelwise and regionwise intra-visit and inter-visit variability measures (STD inter and STD intra) are generally improved with image registration. In the voxelwise analysis, STD inter and STD intra have the same order of magnitude as the $FCRLB_\sigma$. This indicates that the $FCRLB_\sigma$ is a fair estimate of the lowest achievable uncertainty. As for the regionwise analysis, the Monte Carlo estimate of the median $FCRLB_\sigma$ is found to be much lower than the observed variability values in all four scenarios. This may be due to the fact that the assumption that voxels are independent is not satisfied. In a more general perspective, the improvements in regionwise and voxelwise variabilities observed when using image registration techniques could translate into a need for smaller sample sizes in clinical trials.

The uncertainty and reproducibility measures obtained using the complete dataset (4 series) without registration are very similar to those obtained with only one series combined with image registration. This means that a single dataset processed with image registration yields lower uncertainty and as reproducible results as bigger unprocessed datasets. Furthermore, median ADCs are different in these two cases. It has been shown that the DW-MRIs are better aligned when image registration is used, which suggests that registration removes a bias due to misalignments. We therefore discourage the common practice of acquiring multiple image series and averaging the same types of DW-MRIs without applying image registration. This averaging technique indeed requires the acquisition of images that are not necessary if registration is used.

2.6 Conclusion

This chapter shows that image registration can improve the alignment of diffusion-weighted MRIs. Compared to non processed images, registration yields not only different median ADCs, but also ADCs with lower uncertainty and higher reproducibility values. In terms of reproducibility, this is also what Ragheb et al. [109] showed with their local rigid approach, applied to the same datasets. With respect to Gaussian blurring, image registration yields sharper and better aligned images, potentially enabling heterogeneity studies. Furthermore, this chapter shows that it is preferable to consider small but registered datasets rather than big unregistered ones. Indirectly, registration can therefore be considered as a way to reduce acquisition time.

Intravoxel incoherent motion for treatment response monitoring in cystic fibrosis patients with respiratory tract exacerbation

CHAPTER 3

Abstract — Intravoxel incoherent motion (IVIM) is a quantitative imaging technique that yields quantitative parameters linked to molecular-based diffusion (parameter denoted D), blood perfusion (parameter denoted D^*) and volume fraction (parameter denoted f). These parameters are obtained by applying a curve fitting step to diffusion-weighted MR images, in a similar way as done in Chapter 2 for apparent diffusion coefficients (ADC), but using an extended fitting model. In this chapter, our goal is to assess whether the IVIM parameters could be used to monitor treatment response during respiratory tract exacerbation (RTE) in patients with cystic fibrosis (CF). Patients with CF were enrolled and divided into RTE and control groups. Diffusion-weighted MR images were acquired for 11 b-values (0 to 800 s/mm^2) at two time points. RTE patients received antibiotics after baseline image acquisition. Post-processing included motion compensation and a curve fitting step generating the quantitative IVIM parameters D, D^* and f. Statistics were derived from manually delineated volumes of interest. Data comparisons were performed both between the RTE and control groups, and between the baseline and follow-up time points. Besides, correlations with clinical RTE scores were assessed. At baseline, D^* were significantly higher in RTE patients than in controls ($p = 0.008$). For RTE patients, a significant decrease in D^* between baseline and follow-up ($p = 0.036$) was found, while it remained stable in controls. The RTE scores followed a similar behavior ($p < 0.001$). At follow-up, D^* and RTE scores in RTE patients were comparable to their respective distribution in controls. The other quantitative parameters did not evolve in patterns analogous to RTE scores. In this chapter, we show that D^* could be a promising quantitative parameter for capturing clinical response to treatment of patients affected by RTE.

Based upon: J.-M. Guyader, P. Ciet, A. Mazzaro, T. Feiweier, H. A. W. M. Tiddens, G. Morana and S. Klein, "Can intravoxel incoherent motion (IVIM) be used to monitor treatment response in cystic fibrosis (CF) respiratory tract exacerbation (RET)", *submitted*.

3.1 Introduction

Cystic fibrosis (CF) is a lung disease characterised by chronic airways infections, inflammations, respiratory tract exacerbations (RTE) and progressive lung damage [135]. RTE are acute episodes of infection and inflammation, clinically characterised by changes in cough, sputum production, dyspnea, decrease in energy level and appetite, weight loss, and spirometric parameters [46].

Chest X-ray (CXR) and computed tomography (CT) are commonly used for CF monitoring, but both have limitations. CXR only allows the detection of gross lung changes like consolidations, atelectasis and pleural effusions [35]. CT is far more sensitive in localising smaller lung structural changes [84]. However, short-term follow-up with CT is restricted by radiation exposure [69]. This has been the driving force to introduce magnetic resonance imaging (MRI) in CF lung monitoring [156]. Unfortunately, conventional MRI sequences are less sensitive than CT for detecting lung structural changes in CF [23,38]. New sequences have been introduced, but are still at a research stage [34]. The interest for using MRI in the context of CF has also been expanding towards functional imaging [136].

Diffusion-weighted MR imaging (DW-MRI) might have great potential in CF monitoring [7]. So far, it has primarily been used to assess malignancies in thoracic images [86], to identify inflammation [44,97,108,110], or to distinguish between cancer and focal inflammatory lesions [31]. Studies of Ciet et al. [21,22] were among the first to propose the use of DW-MRI to assess lung inflammation. In Ciet et al. [22], semi-quantitative scores [83,130,140] based on DW-MRI were correlated with parameters reflecting clinical symptoms such as the forced expiratory volume (FEV_1), or pulmonary exacerbation RTE scores. In another study, Ciet et al. [21] proposed a fully quantitative approach based on apparent diffusion coefficients (ADC) [74] computed from lung DW-MRI, and correlated them with clinical parameters of disease severity. However, one of the disadvantages of the ADC is that it may incorporate information both about water diffusion and blood perfusion. An alternative is to use intra-voxel incoherent motion (IVIM) [73–75], a technique that can decompose DW-MRI signal in three quantitative parameters: molecular-based diffusion (denoted D), perfusion (denoted D^*) and volume fraction (denoted f). This technique was used for predicting antifungal treatment response in the lung [158], but never yet in a CF context. The present chapter bases itself on the datasets that were used for the ADC study of Ciet et al. [21].

The aim of the present retrospective analysis is to assess whether quantitative IVIM parameters can be used to monitor treatment response during RTE in patients with CF.

3.2 Materials and methods

3.2.1 Research ethics board approval and consent

After institutional review board approval (Protocol no. 314/AULSS9), informed consent was obtained from all patients and/or their guardians. Image acquisition was

performed between 09/2011 and 09/2013.

3.2.2 Study design and participants

This study is a retrospective interpretation of prospectively acquired data, consisting of two visits: baseline and follow-up. The patients with CF taking part to the study were divided into two groups: RTE and controls. All subjects with RTE meeting the Rosenfeld criteria for pulmonary exacerbation [116] were included in the RTE group (see Section 3.2.2.2). CF patients of the control group were consecutively recruited from an annual check-up visit.

3.2.2.1 Inclusion/exclusion criteria

Table 3.1 describes the general inclusion and exclusion criteria, and the RTE scoring criteria capturing clinical parameters related to RTE, used to establish the RTE and control groups. RTE scores, established at baseline and follow-up, are a measure of overall clinical condition related to pulmonary exacerbation.

Table 3.1: *General inclusion and exclusion criteria.*

Inclusion criteria	Exclusion criteria
• Proven CF as evidenced by positive sweat test or gene mutation • Ability to perform reproducible maneuvers with spirometry • Ability to comply with MRI procedures • FEV1 $\geq$ 40 • Age $\geq$ 8 years	• Chronic oxygen therapy • Isolation of Burkholderia cepacia • Lung transplantation • Participation in another trial • Any contraindication to MRI • Pregnancy • Not able or not willing to give consent • Allergic bronchopulmonary aspergillosis

3.2.2.2 Control/exacerbation groups

No universally accepted definition of a respiratory tract exacerbation (RTE) exists. A general definition, described as *clinical need for additional treatment as indicated by acute changes in clinical parameters*, has been recently adopted by the Eurocare CF Working Group. These acute changes during RTE include: change in sputum; increased cough; increased malaise; fatigue or lethargy; anorexia or weight loss; decrease in spirometry outcomes by 10% or more or radiographic changes; and increased dyspnea [136].

These changes are also used by the Rosenfeld criteria to score the clinical likelihood of RTE in patient with CF (Table 3.2). All subjects enrolled in the study were scored with the Rosenfeld criteria for RTE both at baseline at follow-up.

Table 3.2: *Rosenfeld criteria for respiratory tract exacerbation (RTE).*

Rosenfeld criteria	RTE scoring system
• Reduced exercise tolerance • Increased cough • Increased sputum/cough congestion • School or work absenteeism • Increased adventitial sound on lung examination • Reduced appetite • Reduced FEV1 $\geq$ 10 predicted	• Each Rosenfeld criterion is assigned a coefficient • The RTE score is the sum of the coefficients of all criteria (when present) • The threshold value to define RTE is 2.6

3.2.3 Antibiotic treatments

Patients of the RTE group received intravenous antibiotics treatment after the acquisition of the baseline images, according to the presumed infectious pathogen. For the infections due to pseudomonas aeruginosa, a combination of tobramycin (Nebicina, Teofarma, Italy) with another molecule was used, such as ceftazidime (Glazidim, GlaxoSmithKline, UK), cefepime (Maxipime, Bristol-Myers Squibb, USA), meropenem (Merrem, AstraZeneca, UK) or piperacillin/tazobactam (Tazocin, Wyeth Lederle, Pfizer group, USA). For the infections due to staphylococcus aureus one molecule was chosen among oxacillin (Penstapho, Bristol-Myers Squibb, USA), teicoplanin (Targosid, Sanofi-Aventis, France) or vancomycin (Vancomicin Mylan, Mylan, USA). A combination of three antibiotics was used in patients with multiple infections.

Follow-up imaging was performed when each RTE patient was deemed clinically improved by the handling physician. The same median time interval (~20 days) was used for controls.

3.2.4 MRI protocol

DW-MR images were acquired on a 1.5 Tesla MR scanner (MAGNETOM Avanto; Siemens Healthcare, Erlangen, Germany) with a multi-slice 2D single shot echo-planar prototype sequence, in the transverse orientation, for 11 b-values: 0, 10, 20, 30, 50, 70, 100, 150, 200, 400 and 800 s/mm^2. Neither respiratory nor cardiac triggering was used. Across all subjects, the repetition time (TR) was set between 3,800 ms and 6,500 ms (average: 5,051 ms, STD: 663 ms), and the echo time (TE) between 52 and 59 ms (average: 54 ms, STD: 2 ms). Matrix size was set between 160$\times$130 and 180$\times$180, and between 30 and 50 slices (average: 38, STD: 4.5) were acquired with an in-plane spatial resolution varying between 1.78$\times$1.78 mm^2 and 2.60$\times$2.60 mm^2. The slice thickness was set between 6 mm and 7.8 mm (average: 6.34 mm, STD: 0.63 mm). The diffusion gradients were applied along three optimised orthogonal axes.. Each acquisition resulted in 31 images and required between 2.1 and 4.2 minutes (average:

2.9 minutes, STD: 0.4 minute). With this sequence, each DW-MRI was reconstructed from multiple two-dimensional slices acquired in an interleaved manner, as in Chapter 2: the odd slices were acquired first in the inferior–superior direction, followed by the even slices, in the same direction.

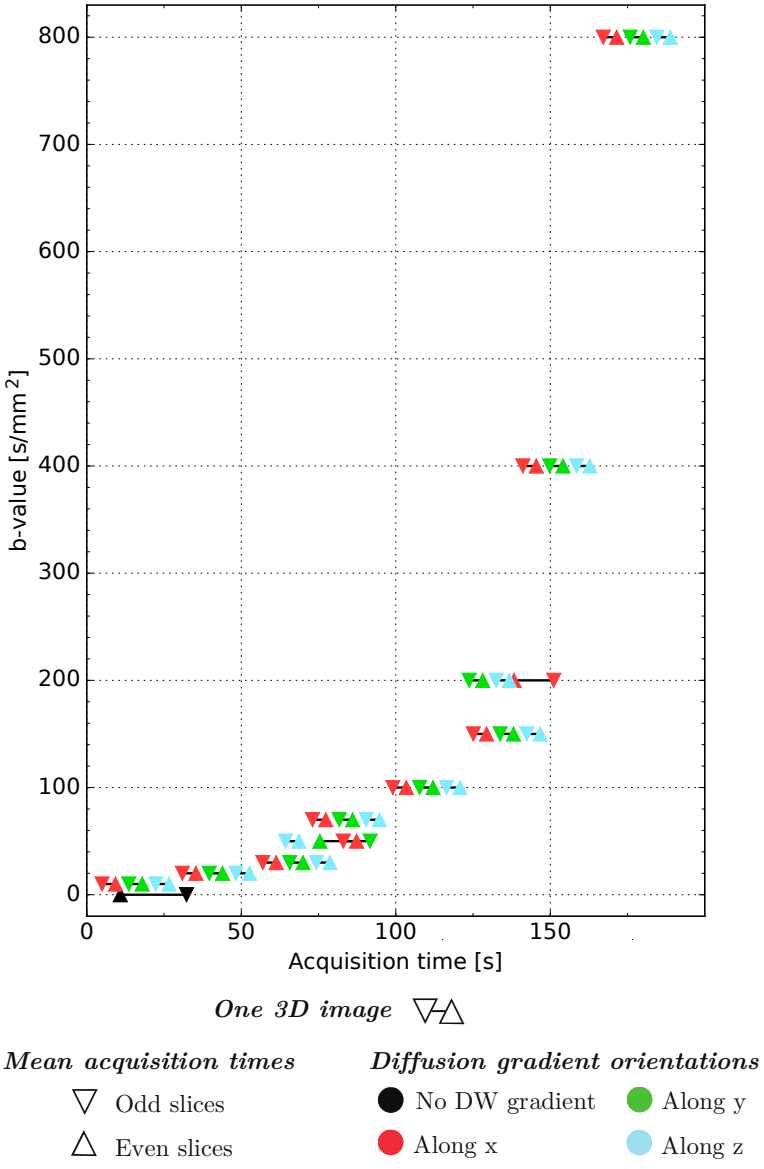


Figure 3.1: *Acquisition sequence. A ▽ symbol represents the mean acquisition time of the odd slices (△ for the even slices).*

All 56 patients considered in the present chapter have been previously reported in Ciet et al. [21]. This prior article proposed a quantitative approach based on ADC, while our approach is based on IVIM. Besides, our approach ensures spatial correspondence between the images using dedicated registration and smoothing techniques.

3.2.5 Image post-processing

The image processing pipeline includes three successive parts: an image registration step, an extraction of IVIM quantitative parameters using curve fitting, and a weighted local smoothing based on Fréchet-Cramér-Rao lower bounds.

3.2.5.1 Step ① – Image registration

The image registration pipeline that we developed in Chapter 2 was designed for interleaved acquisitions such as the one that was used in the context of the present chapter. We therefore applied it to correct for motion-induced artefacts within each DW-MRI (Figure 3.2), and for the misalignments between the DW-MR images. Motion-induced misalignments affecting the DW-MRIs are here mostly due to respiratory motion and patient bulk motion. The registration method that we developed in Chapter 2 compensates for the odd-even slice motion artefacts within each DW-MR images (intra-image registration), and subsequently brings all DW-MR images to the same image space (inter-image registration). Non-rigid transformations based on a three-dimensional transformation model were used [118]. We chose a control point spacing of 64 mm for intra-image registration and 150 mm for inter-image registration. All image registrations were carried out with the open source elastix software [65].

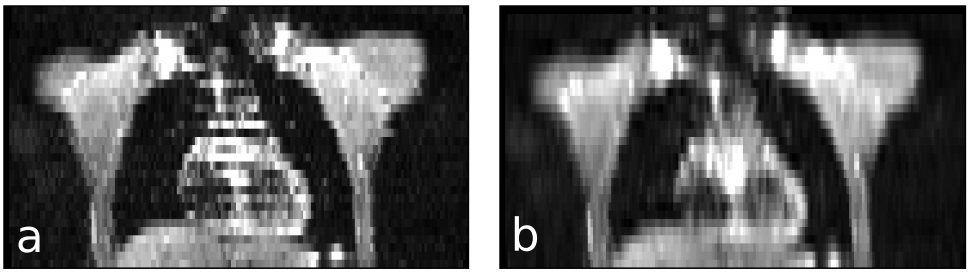


Figure 3.2: *(a) Original DW-MRI coronal view of a patient (b = 0 s/mm²). Note the artefacts between odd and even slices, deriving from the interleaved and free-breathing acquisition. (b) Same DW-MRI after motion compensation.*

3.2.5.2 Step ② – Quantitative IVIM analysis

The subsequent step consisted of extracting quantitative parameters from the registered images. IVIM theory [73–75] indicates that pure molecular diffusion and

microcirculation can be distinguished provided that both multiple low b-values (< 200 s/mm^2) and multiple high b-values (> 200 s/mm^2) are used, which was the case in the present study. IVIM yields quantitative images for molecular-based diffusion (D), blood perfusion (D^*) and volume fraction (f). We extracted the IVIM parameters D, D^* and f from the DW-MR images by applying a two-step curve fitting based on a maximum likelihood estimator.

Le Bihan demonstrated that pure molecular diffusion and microcirculation can be distinguished provided that both multiple low b-values (< 200 s/mm^2) and multiple high b-values (> 200 s/mm^2) are used [85]. These quantitative parameter maps D, D^* and f were obtained by using a curve fitting algorithm based on a maximum likelihood estimator that takes into account the Rician characteristics of the noise in the MR images. The fitting procedure [107] consisted of two parts, the first of which is a mono-exponential fitting based on the following model:

$$S = S_0 \exp(-bD) \tag{3.1}$$

The two parameters for this fitting are S_0 and D. Previous studies showed that micro-circulation and perfusion effects cause deviation from the mono-exponential model for low b-values (i.e., under 50 to 100 s/mm^2) [5, 33, 67, 71–73]. To avoid the influence of perfusion on signal decay, only the diffusion-weighted images with b-values equal to 150, 200, 400 and 800 s/mm^2 were considered for the mono-exponential fitting. The DW-MRIs corresponding to the three different gradient directions were considered as separate measurements in the scalar model presented above. After applying this mono-exponential fitting, the images underwent a more elaborated fitting, based on the following bi-exponential model:

$$S = S_0 \left((1-f)\exp(-bD) + f\exp(-b(D+D^*))\right) \tag{3.2}$$

Similarly to the technique described by Fujima et al. [42], the bi-exponential curve was fitted for S_0, f and D^* while keeping D equal to the value determined in the mono-exponential fitting step. In addition, the fractional volume f was constrained between 0 and 1. Besides, a Fréchet-Cramér-Rao lower bound (FCRLB) was computed at each voxel. The FCRLB provides a lower bound on the variance of the fitted parameters, which allows to quantify the uncertainty of the fitting [107].

3.2.5.3 Step ③ – Weighted local smoothing

The curve fitting is followed by a weighted local smoothing of the quantitative maps D, D^* and f, based on the uncertainty of the fit, aiming to reduce the impact of unreliable voxels on the final statistics. For a given parameter map P among D, D^* or f, the corresponding smoothed version was obtained using the following formula:

$$P_{\text{smoothed}}(\mathbf{x}) = \frac{\sum\limits_{\mathbf{y}} G_\sigma(\mathbf{x},\mathbf{y}) \frac{P(\mathbf{y})}{\text{FCRLB}(\mathbf{y})}}{\sum\limits_{\mathbf{y}} G_\sigma(\mathbf{x},\mathbf{y}) \frac{1}{\text{FCRLB}(\mathbf{y})}} \tag{3.3}$$

with G_σ a Gaussian kernel with $\sigma = 0.5$ voxel, $\mathbf{x}$ the pixel location at which the smoothed version of the parameter map is computed, and $\mathbf{y}$ all possible pixel locations within the support of the kernel G_σ.

3.2.6 Volumes of interest

Quantitative parameters were evaluated in 3D volumes of interests (VOIs) delineated by a thoracic radiologist (P.C.) with 7 years of experience in chest MRI. The presence of a hotspot VOI was determined by comparing the signal of the lung parenchyma with the signal of the spinal cord at the highest b-value (800 s/mm^2). VOIs were delineated at baseline and at follow-up on the $b = 800$ s/mm^2 image with diffusion gradient along z. Zones with a lung parenchyma signal superior or equal to the spinal cord signal were considered as hotspots. Spinal cord served as reference because of its constant high signal due to restricted diffusion, and because it is visible alongside the chest. The baseline VOI was used when the hotspot was not visible at follow-up. Areas with susceptibility artefacts caused by air-tissue interface were excluded from image analysis.

3.2.7 Quantitative parameters

The three IVIM parameters D, D^* and f were considered as quantitative parameters. For a given parameter, the voxel values of each VOI were summarised by taking their average. The fourth quantitative parameter that we considered is the volume (in mm^3) of the VOIs.

3.2.8 Statistical analysis

Analysis of the quantitative parameters was performed at VOI level, treating each VOI as an independent sample. The assumption was made that quantitative parameters extracted from different VOIs are uncorrelated. Parameter distributions between RTE and controls were compared using Mann-Whitney U tests, and between baseline and follow-up using Wilcoxon signed-rank tests. Spearman's rank correlation coefficients were used to measure association between the RTE scores, on the one hand, and the quantitative parameters (D, D^* and f , and the volumes), on the other hand. We considered the p-value significance threshold of 0.05. The p-values reported in the manuscript were adjusted using the Bonferroni correction: we corrected for 4 tests by multiplying the raw p-values by 4. Receiver operating characteristic (ROC) curves were established to assess the area under the curve (AUC), and evaluate the ability of each quantitative parameter to classify patients with or without exacerbation. The parameters were compared to the baseline RTE scores, taken as gold standard for characterising pulmonary exacerbation. Statistical analyses were performed using Python (version 3.5.2), with NumPy (version 1.11.1) and SciPy (version 0.17.1). ROC analyses were performed with SPSS (version 20.0, Chicago, USA).

3.3 Results

3.3.1 Patients' characteristics

Like in the study of Ciet et al. [21], 56 patients with CF were enrolled in the study (22 males, 34 females), with ages varying between 12 and 58 years (average: 24.4 years, STD: 11.2 years). The ages of male patients varied between 12 and 41 years (average: 20.7 years, STD: 8.2 years) and the ages of female patients varied between 12 and 58 (average: 26.8 years, STD: 12.1 years). Among the 56 patients, 29 were included in the RTE group and 27 in the control group. The amount of time between the baseline and follow-up chest MRI varied between 10 and 49 days for patients within the control group (average: 19.6 days, STD: 8.6 days), and between 12 and 77 days for patients within the RTE group (average: 27.7 days, STD: 14.9 days).

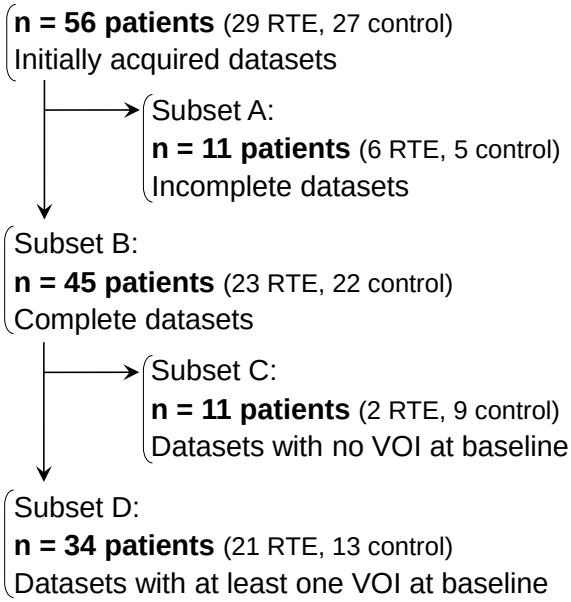


Figure 3.3: *Subsets of patients.*

Not all of the 56 patient datasets were used for analysis (Figure 3.3). Eleven patients had missing images (subset A: 6 patients in the RTE group and 5 in the control group). In 9 out of these 11 cases, the DW-MRIs corresponding to the same b-values had been averaged over the three diffusion gradient directions, which made it impossible to correct for motion artefacts using image registration. The remaining two patients were discarded because the DW-MRIs were not acquired for some b-values. In total, there were 45 patients with correctly acquired datasets (subset B: 23 patients in the RTE group and 22 in the control group). VOI delineation was

Table 3.3: *Number of VOIs delineated for subset D*

	Baseline	Follow-up	
	Manually delinated VOIs	Matching manually delineated VOIs	VOIs transferred from baseline
Control group 13 patients	52	37	15
RTE group 21 patients	89	60	29
All patients 34 patients	141	97	44

performed for these 45 patients. No hotspot VOI was identified at baseline for 11 of these 45 patients (subset C: 2 patients in the RTE group and 9 in the control group). The quantitative analyses were performed on the 34 remaining patients for which at least one VOI was delineated at baseline (subset D: 21 patients in the RTE group and 13 in the control group).

A total number of 238 hotspots VOIs were defined for patients of subset D (Table 3.3). A total of 141 VOIs were manually delineated at baseline. At follow-up, the observer delineated 97 VOIs that could be matched with a VOI at baseline. When no matching hotspot was identified at follow-up, the same VOI as at baseline was used. All subsequent analyses were obtained on subset D.

3.3.2 IVIM parameters and volumes of the VOIs

Examples of fitted quantitative images D, D^* and f are provided in Figure 3.4. Boxplots for the four biomarkers D, D^*, f and the volumes of the VOIs are provided in Figure 3.5. Data from all VOIs (i.e. manually delineated and transferred VOIs) was used to generate the boxplots for parameters D, D^* and f. For the volume parameter, only VOIs that could be matched between baseline and follow-up were considered. The corresponding medians and interquartile ranges (IQR) are presented in Table 3.4.

At baseline, the distribution of D^* and the volumes of the VOIs were significantly higher in the RTE group than in the control group ($p = 0.008$ and $p < 0.001$, respectively). The distributions of D and f were not statistically different between the two groups at baseline ($p = 0.188$ and $p = 0.096$, respectively). At follow-up, the distributions of D, D^* and f were not significantly different between the control and RTE groups ($p = 0.304$, $p = 1.828$, and $p = 0.280$, respectively). The distributions of the volumes remained statistically different at follow-up ($p < 0.001$). In the control group, no significant change was observed for any of the four parameters D, D^*, f and the volumes between baseline and follow-up ($p = 2.492$, $p = 1.020$, $p = 1.756$, and $p = 0.084$, respectively). In the RTE group, significant changes were measured between baseline and follow-up for D^*, f and for the volumes ($p = 0.036$, $p = 0.012$, and $p = 0.008$, respectively), but not for D ($p = 0.748$).

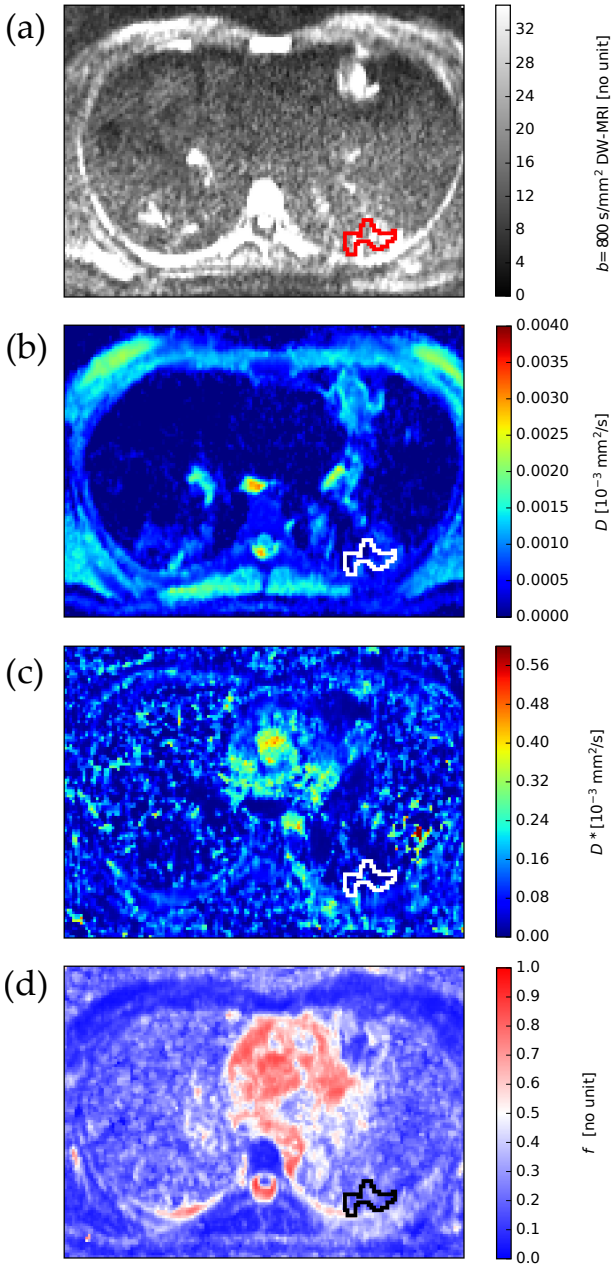


Figure 3.4: *Examples of DW-MRI and fitted results. (a) b = 800 s/mm² DW-MRI image with a VOI in the apical segment of the inferior left lobe, which was thereafter superimposed on the images, (b) D, (c) D^* and (d) f.*

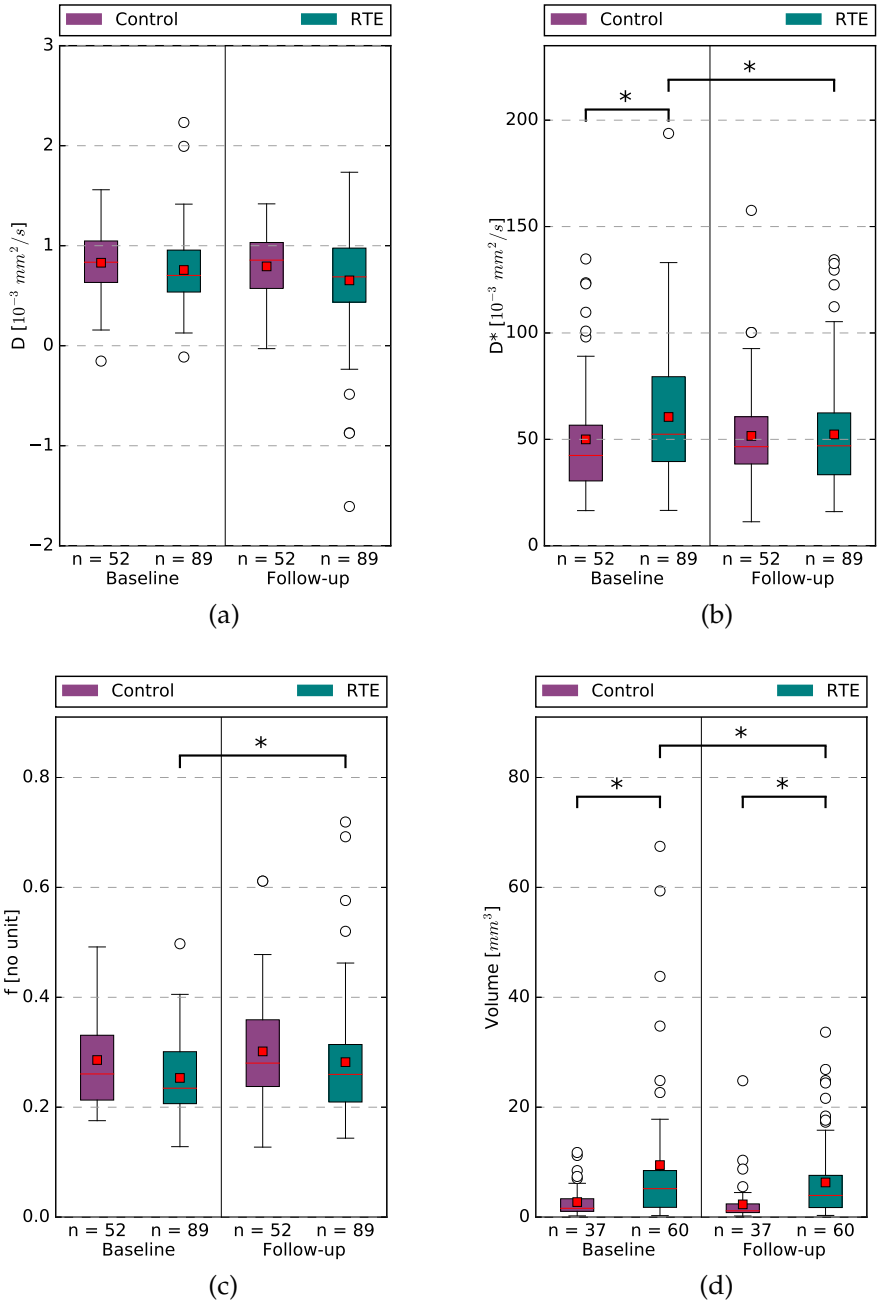


Figure 3.5: *Distributions of (a) D, (b) D^*, (c) f and (d) the volumes of the VOIs. Statistical significance is indicated by *: $p < 0.05$. The mean value for each distribution is indicated by a red square. n indicates the number of VOIs that contribute to the corresponding box-plot.*

Table 3.4: *Median and inter-quartile range (IQR) for the four biomarkers.*

		Baseline		Follow-up	
		Control	RTE	Control	RTE
D	Median	0.835	0.703	0.856	0.689
$[10^{-3}\ \mathrm{mm}^2/\mathrm{s}]$	IQR	0.415	0.420	0.459	0.540
D^*	Median	42.4	52.4	46.6	47.0
$[10^{-3}\ \mathrm{mm}^2/\mathrm{s}]$	IQR	26.2	39.9	22.2	29.1
f	Median	0.261	0.235	0.280	0.260
[*]	IQR	0.118	0.095	0.121	0.105
Volumes	Median	1.88	5.64	1.01	4.80
[mL]	IQR	2.89	9.01	1.79	6.31

3.3.3 RTE scores

Boxplots of RTE scores distributions are provided in Figure 3.6. At baseline, the distributions of RTE scores between the control and RTE groups were significantly different ($p < 0.001$), but no difference was found at follow-up ($p = 1.396$). In the control group, the RTE scores are not statistically different between baseline and follow-up ($p = 0.992$). In the RTE group, the RTE scores were significantly different between baseline and follow-up ($p < 0.001$).

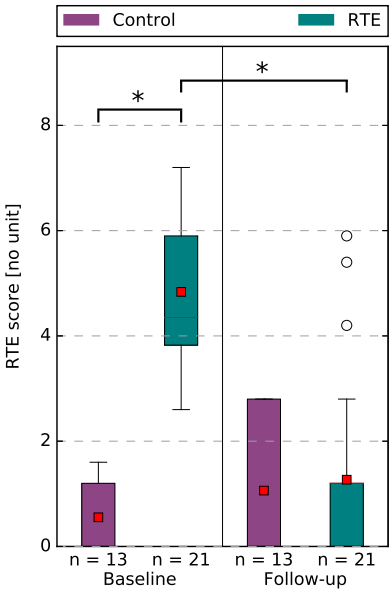


Figure 3.6: *Distributions of RTE scores. Statistical significance is indicated by *: $p < 0.05$. The mean value for each distribution is indicated by a red square. n indicates the number of patients that contribute to the corresponding box-plot.*

3.3.4 Correlations

At baseline, we found a moderate correlation between D^* and the RTE scores ($r = 0.30$, $p < 0.001$), and between the volumes and the RTE scores ($r = 0.33$, $p < 0.001$). At baseline, no significant correlation was found for D ($r = -0.11$, $p = 0.72$) and f ($r = -0.13$, $p = 0.52$). At follow-up, a weak correlation was found only between f and the RTE scores ($r = 0.23$, $p = 0.024$), but no significant correlation was found for D ($r = 0.07$, $p = 1.72$), D^* ($r = 0.14$, $p = 0.44$), and for the volumes ($r = -0.09$, $p = 1.24$).

3.3.5 ROC curves

Accuracy was expressed as area under the curve (AUC). D had a very poor (AUC 0.60, 95% CI 0.50–0.69, $p = 0.056$), $D*$ a poor (AUC 0.64, 95% CI 0.55–0.73, $p = 0.005$), f a poor (AUC 0.62, 95% CI 0.53–0.72, $p = 0.012$) and the volume of the VOIs a poor (AUC 0.67, 95% CI 0.58–0.76, $p = 0.001$) diagnostic accuracy.

3.4 Discussion

This chapter investigates the use of IVIM for treatment response monitoring during RTE. The perfusion D^* and the VOI volumes were significantly higher in RTE patients than in controls, at baseline. This was not the case at follow-up. A similar pattern was observed for the RTE clinical scores. For patients in the RTE group, there was a significant decrease in D^* and RTE scores between baseline and follow-up. Conversely, both D^* and RTE scores remained stable in controls. Additionally, we showed that D^* and RTE scores were correlated at baseline, but not at follow-up.

Lung vascularity is characterised by two components: pulmonary and bronchial. The pulmonary component is dominant in normal conditions. In many inflammatory lung diseases, angiogenesis from bronchial arteries can be prevalent with an increase in the bronchial arteries blood flow from the normal 1% to 35% of cardiac output [151]. Conversely, pulmonary flow decreases due to hypoxic pulmonary vasoconstriction [36]. Wielputz et al. [157] showed a decrease in perfusion score during RTE, which reverted to normal after therapy. In our study, the perfusion D^* was significantly higher in RTE patients than in controls, at baseline, which disappeared at follow-up. This discrepancy between Wielputz's and our results can be explained by the different methods used to evaluate pulmonary perfusion. While contrast-enhanced MR imaging (CEMRI), used by Wielputz, explores perfusion related to pulmonary arteries (which is decreased during inflammation due to hypoxic pulmonary vasoconstriction), IVIM explores perfusion at the microcirculation level, which is increased and mostly represented by bronchial vessels during inflammation.

These results indicate that D^* is a promising non-ionising and non-contrast quantitative imaging biomarker for monitoring inflammation up to complete resolution during RTE, as inflammation and its related hyperperfusion can persist for weeks, even after therapy [3]. IVIM could be used to quantitatively monitor complete inflammation resolution without the use of ionising radiation. Contrary to CEMRI, IVIM presents no risk of gadolinium deposition [47, 98, 134].

VOI volumes distributions followed a pattern analogous to D^*, with an exception: at follow-up, they remained significantly different between both groups, though to a lower extent than at baseline. This suggests that the evolution of lung inflammation is observed more quickly with the quantitative MR parameters D^* than with the hotspots volumes.

A previous study by Ciet et al. [21] extracted semi-quantitative DW-MRI scores and ADCs based on the same DW-MRI dataset used in the present chapter. They obtained ADCs using the same model as for D, but using all b-values, while b-values between 150 and 800 s/mm² were used for D, as recommended in [5,67,71,73]. Ciet et al. [21] showed a significant difference between ADC distributions in the control and RTE groups, at baseline, while we showed no such difference for D. This may be because the ADC incorporates both a diffusion and a perfusion component, while D incorporates diffusion only [72]. This indicates that the most of the inflammatory changes during RTE detected by IVIM are likely perfusion-related, as shown by Wielputz et al. [157].

Investigating the use of IVIM for assessing treatment response to invasive fungal infection (IFI) of the lungs, Yan et al. [158] showed that f was significantly higher in patients with favorable response than in patients with unfavorable response, and inferred that f may be a biomarker for antifungal treatment response. Our study, though focusing on other pathologies and treatments, also indicates that f increased in treated patients between baseline and follow-up. This behavior might be due to hypoxic pulmonary vasoconstriction [36].

Our results indicate that D^*, f, and the VOI volumes have a better diagnostic accuracy than D. However, the diagnostic accuracy of D^*, f and VOI volumes is poor, suggesting that they cannot be used to monitor individual patients. Our results show that these parameters could however be used at a population-based level. These results differ from the results previously reported by Ciet et al. [21] who showed that the ADC had a good diagnostic accuracy, which may be explained by the fact that Ciet et al. [21] selected only the most suspicious hotspot per patient, while we considered all of them.

Our study had several limitations. Firstly, 20% of the datasets were improperly saved and could not be exploited. Secondly, we considered the parameters obtained from the VOIs as independent measurements, even when they came from the same patient. The assumption was made that the VOIs and the parameters extracted from them are uncorrelated, which might not be the case.

3.5 Conclusion

This chapter shows that the IVIM parameter D^*, expressing bronchial artery related perfusion in inflamed tissues, is a promising quantitative parameter sensitive to lung inflammation, and could be used for capturing treatment response in CF patients affected by RTE. In particular, D^* could be used in conjunction with RTE scores, which characterise clinical symptoms only. Compared to the RTE scores, D^* is quantitative biomaker and not a functional qualitative measure. In addition, one of the main advantages of D^* is that it evaluates treatment response per lesion, which is not

possible with RTE scores. Our findings support the potential of IVIM in providing non-contrast, radiation-free quantitative imaging biomarkers for lung inflammation.

Groupwise image registration based on a total correlation dissimilarity measure for quantitative MRI and dynamic imaging data

CHAPTER 4

Abstract — The most widespread technique used to register sets of medical images consists of selecting one image as fixed reference, to which all remaining images are successively registered. This pairwise scheme requires one optimisation procedure per pair of images to register. Pairwise mutual information is a common dissimilarity measure applied to a large variety of datasets. Alternative methods, called groupwise registrations, have been presented to register two or more images in a single optimisation procedure, without the need of a reference image. Given the success of mutual information in pairwise registration, we adapt one of its multivariate versions, called total correlation, in a groupwise context. We justify the choice of total correlation among other multivariate versions of mutual information, and provide full implementation details. The resulting total correlation measure is remarkably close to measures previously proposed by Huizinga et al. based on principal component analysis. Our experiments, performed on five quantitative imaging datasets and on a dynamic CT imaging dataset, show that total correlation yields registration results that are comparable to Huizinga's methods. Total correlation has the advantage of being theoretically justified, while the measures of Huizinga et al. were designed empirically. Additionally, total correlation offers an alternative to pairwise mutual information on quantitative imaging datasets.

Based upon: J.-M. Guyader, W. Huizinga, D. H. J. Poot, M. van Kranenburg, A. Uitterdijk, W. J. Niessen and S. Klein, "Groupwise image registration based on a total correlation dissimilarity measure", *Scientific Reports*, vol. 8, pp. 13112, 2018.

4.1 Introduction

Intensity-based image registration using the maximisation of mutual information is commonly used for aligning pairs of medical images that do not have similar intensity distributions, or are acquired from different modalities [87, 105, 149]. Mutual information belongs to the family of pairwise dissimilarity measures. Pairwise methods quantify the alignment of a moving image with a fixed reference image. The optimisation process performed in the context of pairwise registration therefore considers only two images simultaneously.

Nowadays, imaging datasets often contain more than two images, acquired from different modalities, different time points or different subjects, for instance. When more than two images have to be registered, the common pairwise registration scheme using a fixed reference image (Section 1.4.1) is not the most adapted scheme. Firstly, the choice of reference image to which the remaining image are registered can be arbitrary, but may also influence the registration results, as shown by Geng et al. [45]. Secondly, the pairwise nature of that registration scheme does not allow the registration of all images in a single optimisation procedure, which prevents taking into account all image information simultaneously.

In this chapter, we focus on symmetric groupwise image registration. This means that the registration of two or more images is performed in a single optimisation procedure, with the additional condition that all images should play an analogous role during registration (i.e. no image should be taken as fixed reference, for instance). Given the success of the mutual information dissimilarity measure in pairwise image registration, this chapter specifically focuses on dissimilarity measures that could extend the use of mutual information concepts to groupwise registration.

Though the formulation of mutual information for two images is unique, several multivariate versions exist for its generalisation for more than two images. These dissimilarity measures are called interaction information [92], total correlation [153] and dual total correlation [51]. Total correlation is the groupwise dissimilarity measure that we propose to adapt in the context of groupwise image registration.

Competing state-of-the-art dissimilarity measures for groupwise registration include the sum of variances developed by Metz et al. [94], the groupwise mutual information method of Bhatia et al. [14], and the groupwise dissimilarity measures based on principal component analysis (PCA) previously developed by Huizinga et al. [55]. The expression of the total correlation dissimilarity measure that we propose is remarkably close to Huizinga's PCA-based groupwise dissimilarity measures, which were shown to outperform competing pairwise and groupwise state-of-the-art methods on qMRI datasets. The experiments conducted in this chapter consist of using groupwise total correlation for the registration of a dynamic CT imaging dataset, and of five quantitative magnetic resonance imaging (qMRI) image datasets. Registration results are compared to Huizinga's methods, but also to pairwise registration based on mutual information.

4.2 Materials and methods

Let us consider $\mathcal{M} = \{M_1, ..., M_G\}$, a series of G images that have to be registered. Each image M_g, consists of N voxels. To quantify how well the G images are aligned, a dissimilarity measure has to be defined. In this chapter, we consider dissimilarity measures based on the concepts of mutual information. We choose the convention to formulate the measures as dissimilarity measures instead of similarity measures, so that the registration problem can be written as a cost function minimisation problem.

4.2.1 Pairwise mutual information

Mutual information is a robust measure that is commonly used for the pairwise registration of datasets of medical images, including multimodal datasets [105]. For $G = 2$ images M_1 and M_2, negated mutual information $\mathcal{D}_{\mathrm{MI}}$ is computed as follows [105, 149]:

$$\mathcal{D}_{\mathrm{MI}}(M_1, M_2) = H(M_1, M_2) - H(M_1) - H(M_2) \tag{4.1}$$

with $H(M_1)$ the entropy [121] of image M_1, $H(M_2)$ the entropy of image M_2, and $H(M_1, M_2)$ the joint entropy of M_1 and M_2. For two images M_1 and M_2, the joint entropy can be computed as follows [25]:

$$H(M_1, M_2) = -\sum_{x_1}\sum_{x_2} P(x_1, x_2) \ln\left[P(x_1, x_2)\right] \tag{4.2}$$

where x_1 and x_2 represent the discrete values of images M_1 and M_2, respectively. $P(x_1, ...x_n)$ is the probability of these values occurring together. $P(x_1, x_2) \ln\left[P(x_1, x_2)\right]$ is defined to be 0 if $P(x_1, x_2)$ equals 0.

When the dataset of images to register contains $G > 2$ images, it is still possible to use a pairwise method to register the images, but several independent registration procedures have to be performed. A typical method consists of selecting one of the images as fixed reference, and then successively applying pairwise registration with the remaining $G - 1$ images considered as moving images (Figure 1.6a, Section 1.4.2). This scheme is not well suited to registration problems for which there is no obvious reference image. Besides, the registration results may be different according to the choice of fixed reference image, as shown by Geng et al. [45]. Seghers et al. [120] introduced a method that we will refer to as semi-groupwise, which is based on multiple pairwise registrations and does not require the selection of a reference space. For each i, M_i is taken as fixed image and $G - 1$ independent registration are performed between each remaining image, M_j, yielding $G - 1$ transformations $T_{i\to j}$ per fixed image M_i. Each image M_i is then resampled into an average or mid-point image space using $\bar{T}_i^{-1}(\boldsymbol{x})$, the inverse of the arithmetic mean of the transformations $T_{i\to j}$ (Figure 1.6b). The method of Seghers et al. [120] has the disadvantage of requiring $G \times (G - 1)$ registration procedures, which becomes computationally complex when G grows. It also does not allow to register all images in a single optimisation procedure.

4.2.2 Groupwise dissimilarity measures based on multivariate mutual information

Groupwise registration techniques allow to register $G \geq 2$ images in one optimisation procedure (Figure 1.6c, Section 1.4.2). In this chapter, we will focus on groupwise techniques that treat the images equally. In particular, the order in which the images are supplied should have no influence on the value of the groupwise dissimilarity measure $\mathcal{D}(M_1, M_2, ..., M_G)$, and therefore no influence on the registration results.

This chapter more precisely focuses on groupwise generalisations of mutual information, given the wide interest and range of applications of that dissimilarity measure in the context of pairwise image registration [105]. There exist multiple multivariate forms of mutual information [51, 92, 153], the concepts of which can be used for groupwise image registration.

The first multivariate generalisation of mutual information is known as interaction information [92], denoted $\mathcal{D}_{\mathrm{II}}$. It measures the amount of information shared by all the images. For the G images of $\mathcal{M} = \{M_1, ..., M_G\}$, the negated interaction information is written as follows:

$$\mathcal{D}_{\mathrm{II}}(\mathcal{M}) = \sum_{V \subseteq \mathcal{M}} (-1)^{G-|V|} H(V) \tag{4.3}$$

with V any subset of images of $\mathcal{M}$, $|V|$ the number of images in the corresponding subset, and $H(V)$ the joint entropy of the subset V. For G images $M_1 ... M_G$ the joint entropy can be computed as follows:

$$H(M_1, ..., M_G) = -\sum_{x_1} ... \sum_{x_G} P(x_1, ..., x_G) \ln \left[P(x_1, ..., x_G)\right] \tag{4.4}$$

where the $x_1, ..., x_G$ are the values of images $M_1, ..., M_G$, respectively. The same definitions as for $P(x_1, x_2)$ and $P(x_1, x_2) \ln \left[P(x_1, x_2)\right]$ are directly extended for $P(x_1, ..., x_G)$ and$P(x_1, ..., x_G) \ln \left[P(x_1, ..., x_G)\right]$. Interaction information quantifies the information shared together by images $M_1, ..., M_G$ [11]. This means that if at least one of the images of $\mathcal{M}$ shares no information with all other images, the interaction information will be zero [11, 43].

The second form of multivariate mutual information, called total correlation [153], measures the amount of information shared between any subset of $\mathcal{M}$. The negated total correlation is written as:

$$\mathcal{D}_{\mathrm{TC}}(\mathcal{M}) = H(\mathcal{M}) - \left[\sum_{g=1}^{G} H(M_g)\right] \tag{4.5}$$

with $H(\mathcal{M})$ the joint entropy of the images of the set $\mathcal{M} = \{M_1, ..., M_G\}$.

The third form is a refinement of total correlation called dual total correlation [51]. The negated dual total correlation is formulated as follows:

$$\mathcal{D}_{\mathrm{DTC}}(\mathcal{M}) = \left[\sum_{g=1}^{G} H(M_g | (\mathcal{M} \setminus M_g))\right] - H(\mathcal{M}) \tag{4.6}$$

with $\mathcal{M}\backslash M_g$ the set of images $\{M_1, ..., M_G\}$ with M_g removed. $H(M_g|(\mathcal{M}\backslash M_g))$ is the conditional entropy [25] of M_g given $\mathcal{M}\backslash M_g$. In other terms, this means that $H(M_g|(\mathcal{M}\backslash M_g))$ is the entropy of the image M_g given that the knowledge of images $\{M_1, ..., M_{g-1}, M_{g+1}, ..., M_G\}$ is known.

Theoretically, both total correlation and dual total correlation quantify the amount of shared information between all possible combinations of images, while interaction information only quantifies the amount of information shared by all images [138]. Venn diagrams [25, 138, 143] for $\mathcal{D}_{\mathrm{II}}$, $\mathcal{D}_{\mathrm{TC}}$ and $\mathcal{D}_{\mathrm{DTC}}$ are shown in Figure 4.1. In the context of image registration, $\mathcal{D}_{\mathrm{TC}}$ and $\mathcal{D}_{\mathrm{DTC}}$ seem more adapted than $\mathcal{D}_{\mathrm{II}}$ in the sense that they are built to quantify shared information not only between all images, but also between any of their subsets [11, 43]. In particular, including an image with little dependence towards the others would impair the registration of the remaining images when using $\mathcal{D}_{\mathrm{II}}$, while this would theoretically not be the case when using $\mathcal{D}_{\mathrm{TC}}$ or $\mathcal{D}_{\mathrm{DTC}}$. We therefore chose to consider the dissimilarity measures based on total correlation to adapt a multivariate version of mutual information as a groupwise dissimilarity measure.

4.2.3 Groupwise total correlation

In this section, we describe how total correlation, as expressed in Formula (4.5), can be brought to practical use in the context of image registration. As such, computing total correlation implies computing the joint entropy $H(\mathcal{M})$, but this computation is subject to the curse of dimensionality [12]: the evaluation of joint entropy requires to compute a G-dimensional joint histogram that becomes increasingly sparser as G increases, and therefore becomes computationally prohibitive.

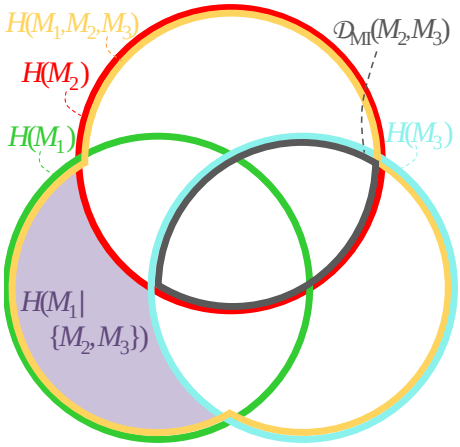
Let us consider a random variable $\mathbf{X} \in \mathbb{R}^G$ following a G-variate normal distribution given by:

$$f(\mathbf{X}) = \frac{1}{\sqrt{\det(2\pi\mathbf{C})}} \exp\left(-\frac{1}{2}(\mathbf{X}-\mu)^{\mathrm{T}}\,\mathbf{C}^{-1}(\mathbf{X}-\mu)\right) \tag{4.7}$$

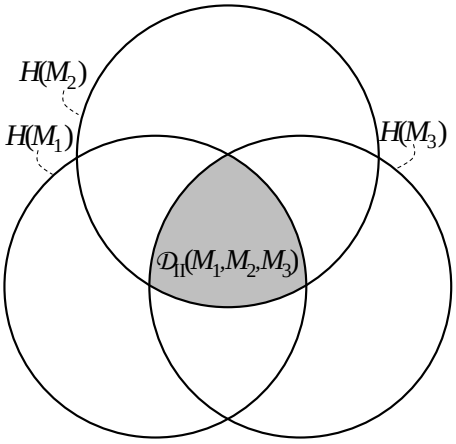
with $\boldsymbol{\mu} \in \mathbb{R}^G$ an expectation vector, $\mathbf{C} \in \mathbb{R}^{G\times G}$ a covariance matrix, and with $\det(\,.\,)$ the determinant operator. Ali Ahmed et al. [1] have shown that the entropy of the multivariate normal variable $\mathbf{X}$ may be written as:

$$H(\mathbf{X}) = \frac{G}{2} + \frac{G}{2}\ln(2\pi) + \frac{1}{2}\ln(\det(\mathbf{C})) \tag{4.8}$$

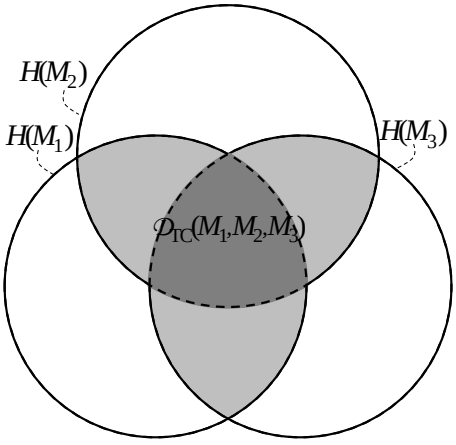
To circumvent the curse of dimensionality, and make it possible to use registration in a groupwise manner on datasets containing any number $G \geq 2$ images, we propose to use Equation (4.8) in the context of G images $\mathcal{M} = \{M_1, ..., M_G\}$. For the sake of efficient calculation of the entropy, we approximate the intensity distribution of the images by a joint normal distribution, and we make the hypothesis that the minimum of the resulting cost function is sill a good solution for the underlying registration problem. Let $\mathbf{M}$ be a $N \times G$ matrix in which each image M_g is represented



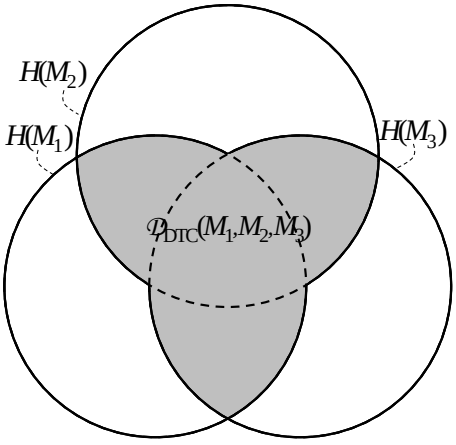
(a) Entropies and mutual information $\mathcal{D}_{\mathrm{MI}}$



(b) Interaction information $\mathcal{D}_{\mathrm{II}}$



(c) Total correlation $\mathcal{D}_{\mathrm{TC}}$



(d) Dual total correlation $\mathcal{D}_{\mathrm{DTC}}$

Figure 4.1: *Venn diagram representations of the entropies, mutual information and multivariate versions of mutual information, for three images M_1, M_2 and M_3. (a) The green, red and cyan circle represent the entropy of each image. The fact that the images share information is symbolised by overlapping circles. The joint entropy $H(M_1, M_2, M_3)$ quantifies the amount of information brought about by the three images, and is symbolised by the union of the three circles (in orange). The mutual information $\mathcal{D}_{\mathrm{MI}}(M_2, M_3)$ of two images M_2 and M_3 is the information that these two images share, and is symbolised by the intersection of the two circles representing the entropies of images M_2 and M_3 (grey circle). Subfigures (b), (c) and (d) were constructed based on Equations* (4.3), (4.5) *and* (4.6)*. In (c), the light grey areas indicate the corresponding contributions to the dissimilarity measure are twice less than the dark grey areas.*

by a column. The matrix $\mathbf{C}$ of covariances between the images M_g is obtained as follows:

$$\mathbf{C} = \frac{1}{N-1}\left(\mathbf{M} - \overline{\mathbf{M}}\right)^{\mathrm{T}}\left(\mathbf{M} - \overline{\mathbf{M}}\right) \tag{4.9}$$

with $\overline{\mathbf{M}}$, a matrix that has in each of its column the column-wise average of $\mathbf{M}$. To make the method robust to linear intensity scalings and offsets, we incorporate an intensity standardisation (i.e. z-score) within the definition of the dissimilarity measure. This is done by computing the entropy $H(\mathcal{M})$ using the correlation matrix $\mathbf{K}$ instead of the covariance matrix $\mathbf{C}$, with:

$$\mathbf{K} = \mathbf{\Sigma}^{-1}\mathbf{C}\mathbf{\Sigma}^{-1} \tag{4.10}$$

with $\mathbf{\Sigma}$ a diagonal matrix with the standard deviations of the columns of $\mathbf{M}$ as its diagonal elements. A diagonal element Σ_{gg} of Σ verifies:

$$\Sigma_{gg} = \frac{1}{N-1}\sum_{i=1}^{N}(M_{g,i} - \overline{M}_g)^2 \tag{4.11}$$

where the $M_{g,i}$ are the individual voxel values and $\overline{M}_g$ the average voxel value of image $\overline{M}_g$. By construction, each diagonal element of the correlation matrix $\mathbf{K}$ is equal to 1.

The expression of the joint entropy therefore becomes:

$$H(\mathcal{M}) = \frac{G}{2} + \frac{G}{2}\ln(2\pi) + \frac{1}{2}\ln(\det(\mathbf{K})) \tag{4.12}$$

Equation (4.12) can also be used to derive the marginal entropies $H(M_g)$. When considering only one image M_g, the correlation matrix $\mathbf{K}$ is the scalar 1. All $H(M_g)$ are therefore constant and equal to:

$$H(M_g) = \frac{1}{2} + \frac{1}{2}\ln(2\pi) \tag{4.13}$$

By combining Equations (4.5), (4.12) and (4.13), we define the dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ based on total correlation as follows:

$$\mathcal{D}_{\mathrm{TC}}(\mathcal{M}) = \frac{1}{2}\ln(\det(\mathbf{K}))) = \frac{1}{2}\sum_{j=1}^{G}\ln\lambda_j \tag{4.14}$$

using $\det(\mathbf{K}) = \prod_{j=1}^{G}\lambda_j$, with λ_j the j^{th} eigenvalue of $\mathbf{K}$, and $\lambda_j > \lambda_{j+1}$. Such a simple expression was not found for dual total correlation, which is why we selected total correlation as groupwise dissimilarity measure.

4.2.4 Gradient-based optimisation and implementation

To implement the approximated version of total correlation $\mathcal{D}_{\mathrm{TC}}$ provided in Equation (4.14), we define an interpolation scheme based on B-splines. This scheme associates with each original image M_g a continuous and differentiable function $M_g(\boldsymbol{x})$ of the spatial coordinate $\boldsymbol{x}$. The aim is to simultaneously bring the images $M_g(\boldsymbol{x})$ to an average space by means of a transformation $\boldsymbol{T}(\boldsymbol{x}, \boldsymbol{\mu})$, where $\boldsymbol{\mu}$ is a vector containing the parameters $\boldsymbol{\mu}_g$ that correspond to the transformation $\boldsymbol{T}_g(\boldsymbol{x}, \boldsymbol{\mu}_g)$ related to each image M_g. Examples of transformation models are the affine model, or the non-rigid model in which deformations are modeled by cubic B-splines [118].

In the groupwise scheme, the measure $\mathcal{D}$ quantifies the dissimilarity between all transformed images $M_g(\boldsymbol{T}_g(\boldsymbol{x}, \boldsymbol{\mu}_g))$. Groupwise registration can therefore be formulated as the constrained minimisation of the dissimilarity measure $\mathcal{D}$ with respect to $\boldsymbol{\mu}$, as previously proposed by Huizinga et al. [55]:

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{D}(M_1(\boldsymbol{T}_1(\boldsymbol{x}, \boldsymbol{\mu}_1)), ..., M_G(\boldsymbol{T}_G(\boldsymbol{x}, \boldsymbol{\mu}_G))) \tag{4.15}$$

subject to the following constraint, allowing to define a mid-point space [10].

$$\sum_{g=1}^{G} \boldsymbol{\mu}_g = \mathbf{0} \tag{4.16}$$

The implementation of the total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ was performed as part of the open source software package elastix [65]. The adaptive stochastic gradient descent (ASGD) developed by Klein et al. [64] is used as optimisation method for image registration. This method randomly samples positions in the image space at each iteration in order to reduce computation time. Sampling is done off the voxel grid, which was shown to be necessary to reduce interpolation artefacts [65]. A multi-resolution strategy is used: the images are Gaussian-blurred with a certain standard deviation, which is decreased at each resolution level. This means that the large deformations are corrected first, and that finer deformations are corrected in subsequent levels. Linear interpolation is used to interpolate the images during registration, which reduces computation time, but cubic B-spline interpolation was used to produce the final registered images. For the chosen ASGD optimisation method, the gradient of the dissimilarity measure is needed. Based on Equation (4.14) and van der Aa et al. [142], we find:

$$\frac{\partial \mathcal{D}_{\mathrm{TC}}}{\partial \boldsymbol{\mu}} = \frac{1}{2} \sum_{j=1}^{G} \frac{1}{\lambda_j} \frac{\partial \lambda_j}{\partial \boldsymbol{\mu}} = \frac{1}{2} \sum_{j=1}^{G} \frac{1}{\lambda_j} \left(\boldsymbol{v}_j^T \frac{\partial \mathbf{K}}{\partial \boldsymbol{\mu}} \boldsymbol{v}_j \right) \tag{4.17}$$

where $\boldsymbol{v}_j^T$ is the j^{th} eigenvector of $\mathbf{K}$. Similarly to van der Aa et al. [142], we make the assumption that the repetition of eigenvalues is unlikely.

When the eigenvalues λ_j tend towards zero, evaluating $\mathcal{D}_{\mathrm{TC}}$ implies taking the natural logarithm of a near-zero number (as shown in Equation (4.14)), which might result in a failing optimisation. We therefore introduce an adjusting constant $c \in \mathbb{R}^+$ that is added to the eigenvalue λ_j before taking the natural logarithm:

$$\mathcal{D}_{\mathrm{TC}}(\mathcal{M}) = \frac{1}{2}\ln(\det(\mathbf{K} + c\mathbf{I})) = \frac{1}{2}\sum_{j=1}^{G}\ln(\lambda_j + c) \tag{4.18}$$

where $\mathbf{I}$ is the identity matrix. The gradient of the adjusted total correlation dissimilarity measure therefore becomes:

$$\frac{\partial \mathcal{D}_{\mathrm{TC}}}{\partial \boldsymbol{\mu}} = \frac{1}{2}\sum_{j=1}^{G}\frac{1}{\lambda_j + c}\frac{\partial \lambda_j}{\partial \boldsymbol{\mu}} = \frac{1}{2}\sum_{j=1}^{G}\frac{1}{\lambda_j + c}\left(\boldsymbol{v}_j^T \frac{\partial \mathbf{K}}{\partial \boldsymbol{\mu}}\boldsymbol{v}_j\right) \tag{4.19}$$

To derive an appropriate value for c, we make the assumption that the first mode, corresponding to the eigenvalue λ_1, accounts for half of the total data variation. Given that the trace of $\mathbf{K}$ is equal to the sum of its eigenvalues, we can write that $\mathrm{tr}(\mathbf{K}) = \sum_{i=1}^{G}\lambda_i$. In addition, the diagonal elements of the correlation matrix $\mathbf{K}$ are all equal to 1, which induces that $\mathrm{tr}(\mathbf{K}) = G = \sum_{i=1}^{G}\lambda_i$. The assumption that the first mode accounts for half of the total data variation therefore yields $\lambda_1 = G/2$. We then constrain the ratio $(\lambda_1 + c)/(\lambda_G + c)$ to G, so that the weights $1/(\lambda_\mathrm{i} + c)$ in Equation (4.19) remain within a known, finite range. We also make the assumptions that $c << G$ and that $\lambda_\mathrm{G} << c$. This leads to the solution $c = 0.5$. In addition to solving a computational issue, the constant c introduces a lower bound on the variance associated with each eigenvector. Initial experiments confirmed that with this choice for c, occasional numerical instabilities were successfully eliminated, while not visibly affecting the results in other cases.

Based on Equation (4.10), one can derive the following expression for $\partial\mathbf{K}/\partial\mu_p$:

$$\begin{aligned}
\frac{\partial \mathbf{K}}{\partial \mu_p} =& \frac{\partial}{\partial \mu_p}\left(\frac{1}{N-1}\boldsymbol{\Sigma}^{-1}\left(\mathbf{M} - \overline{\mathbf{M}}\right)^{\mathrm{T}}\left(\mathbf{M} - \overline{\mathbf{M}}\right)\boldsymbol{\Sigma}^{-1}\right) \\
=& \frac{1}{N-1}\left[\frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \mu_p}\left(\mathbf{M} - \overline{\mathbf{M}}\right)^{\mathrm{T}}\left(\mathbf{M} - \overline{\mathbf{M}}\right)\boldsymbol{\Sigma}^{-1}\right. \\
&+ \boldsymbol{\Sigma}^{-1}\left(\frac{\partial \mathbf{M}}{\partial \mu_p} - \frac{\partial \overline{\mathbf{M}}}{\partial \mu_p}\right)^{\mathrm{T}}\left(\mathbf{M} - \overline{\mathbf{M}}\right)\boldsymbol{\Sigma}^{-1} \\
&+ \boldsymbol{\Sigma}^{-1}\left(\mathbf{M} - \overline{\mathbf{M}}\right)^{\mathrm{T}}\left(\frac{\partial \mathbf{M}}{\partial \mu_p} - \frac{\partial \overline{\mathbf{M}}}{\partial \mu_p}\right)\boldsymbol{\Sigma}^{-1} \\
&\left.+ \boldsymbol{\Sigma}^{-1}\left(\mathbf{M} - \overline{\mathbf{M}}\right)^{\mathrm{T}}\left(\mathbf{M} - \overline{\mathbf{M}}\right)\frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \mu_p}\right]
\end{aligned} \tag{4.20}$$

The property of commutativity of the dot product yields:

$$\boldsymbol{v}^{\mathrm{T}}\mathbf{A}\mathbf{B}\boldsymbol{v} = \boldsymbol{v}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\boldsymbol{v} \tag{4.21}$$

with $\mathbf{A}$ and $\mathbf{B}$, two matrices and $\boldsymbol{v}$ a vector. Using Equations (4.19), (4.20), (4.21), and the fact that $\boldsymbol{\Sigma}^{-1}$ is a symmetric matrix, the derivative of $\mathcal{D}_{\mathrm{TC}}$ with respect to an element μ_p becomes:

$$\begin{aligned}\frac{\partial \mathcal{D}_{\mathrm{TC}}}{\partial \mu_p} =& \frac{1}{N-1}\sum_{j=1}^{G}\Bigg[\frac{1}{\lambda_j + c} \\ &\times \Bigg\{ \boldsymbol{v}_j^T \boldsymbol{\Sigma}^{-1} \left(\mathbf{M}-\overline{\mathbf{M}}\right)^{\mathrm{T}} \left(\mathbf{M}-\overline{\mathbf{M}}\right) \frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \mu_p} \boldsymbol{v}_j \\ &+ \boldsymbol{v}_j^T \boldsymbol{\Sigma}^{-1} \left(\mathbf{M}-\overline{\mathbf{M}}\right)^{\mathrm{T}} \left(\frac{\partial \mathbf{M}}{\partial \mu_p} - \frac{\partial \overline{\mathbf{M}}}{\partial \mu_p}\right) \boldsymbol{\Sigma}^{-1} \boldsymbol{v}_j \Bigg\}\Bigg]\end{aligned} \tag{4.22}$$

To obtain $\partial \boldsymbol{\Sigma}^{-1}/\partial \mu_p$, the diagonal elements $\boldsymbol{\Sigma}_{gg}^{-1}$ of the diagonal matrix $\boldsymbol{\Sigma}^{-1}$ can be derived one by one:

$$\begin{aligned}\frac{\partial \Sigma_{gg}^{-1}}{\partial \mu_p} =& \frac{\partial}{\partial \mu_p}\left(\frac{1}{N-1}\sum_{i=1}^{N}(M_{ig}-\overline{M}_g)^2\right)^{-\frac{1}{2}} \\ =& -\frac{\Sigma_{gg}^{-3}}{N-1}\left[\left(\mathbf{M}-\overline{\mathbf{M}}\right)^{\mathrm{T}}\left(\frac{\partial \mathbf{M}}{\partial \mu_p} - \frac{\partial \overline{\mathbf{M}}}{\partial \mu_p}\right)\right]_{gg}\end{aligned} \tag{4.23}$$

The quantity $\partial \mathbf{M}/\partial \mu_p$ is computed as follows:

$$\frac{\partial M_g(\boldsymbol{T}_g(\boldsymbol{x}, \boldsymbol{\mu}_g))}{\partial \mu_p} = \left(\frac{\partial M_g}{\partial \boldsymbol{x}}\right)^{\mathrm{T}}_{\boldsymbol{T}_g(\boldsymbol{x}, \boldsymbol{\mu}_g)} \left(\frac{\partial \boldsymbol{T}_g}{\partial \mu_p}\right)_{(\boldsymbol{x}, \boldsymbol{\mu}_g)} \tag{4.24}$$

It was verified that the derivative $\partial \overline{\mathbf{M}}/\partial \mu_p$ of the mean intensities was negligibly small and were therefore be ignored in the actual implementation.

4.2.5 Related groupwise dissimilarity measures

Huizinga et al. [55] previously presented two dissimilarity measures, the expressions of which are close to the total correlation measure presented in this chapter (Equation (4.18)). Huizinga's dissimilarity measures are based on principal component analysis (PCA), and were originally designed for the registration of multi-parametric datasets of images, i.e. datasets $\{M_1, ..., M_G\}$ for which the images M_g are characterised by an underlying model m_g describing their intensity values, such that:

$$M_g(\boldsymbol{x}) = m_g(\boldsymbol{\theta}(\boldsymbol{x})) + \epsilon(\boldsymbol{x}) \tag{4.25}$$

with $\boldsymbol{\theta}$ a vector containing the parameters of the model, and ϵ the noise at coordinate $\boldsymbol{x}$. An example of such model is the monoexponential model $m_g(\boldsymbol{\theta}) = S_0 \exp(-b_g \boldsymbol{u}_g^{\mathrm{T}} \mathbf{D} \boldsymbol{u}_g)$ used in diffusion tensor imaging, with $\boldsymbol{\theta} = (S_0, D_{11}, D_{12}, D_{13}, D_{22}, D_{23}, D_{33})$, $\boldsymbol{u}_g$ the diffusion gradient direction vector, $\mathbf{D}$ a 3×3 symmetric diffusion tensor, and b the b-value [74].

Huizinga's dissimilarity measures rely on the idea that an aligned set of multi-parametric images can be described by a small number of high eigenvalues, since the underlying model m_g is low-dimensional, i.e. the size Γ of $\boldsymbol{\theta}$ is lower than G. A misaligned set of multi-parametric images would, on the contrary, be characterised by a flatter eigenvalue spectrum: more eigenvalues of average intensity are required for

describing the data in that case. The first dissimilarity measure introduced by Huizinga et al. [55], denoted $\mathcal{D}_{\mathrm{PCA}}$, is the difference between the sum of all eigenvalues and the sum of the first few eigenvalues:

$$\mathcal{D}_{\mathrm{PCA}}(\mathcal{M}) = \sum_{j=1}^{G} \lambda_j - \sum_{j=1}^{L} \lambda_j = \sum_{j=L+1}^{G} \lambda_j \tag{4.26}$$

with L a user-defined constant with $1 \leq L \leq G$, and $\sum_{j=1}^{G} \lambda_j = \mathrm{tr}(\mathbf{K}) = G$. This means that $\mathcal{D}_{\mathrm{PCA}}$ is the sum of the lowest $G - L$ eigenvalues. Contrary to $\mathcal{D}_{\mathrm{PCA}}$, the second dissimilarity measure designed by Huizinga et al., denoted $\mathcal{D}_{\mathrm{PCA2}}$, does not require the selection of an arbitrary cut-off L. It consists of weighting the last eigenvalues more than the first ones:

$$\mathcal{D}_{\mathrm{PCA2}}(\mathcal{M}) = \sum_{j=1}^{G} j\lambda_j \tag{4.27}$$

$\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$ were developed based on the concepts of PCA, while the total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ presented in this chapter is a multivariate derivation of mutual information. Nevertheless, the expressions of $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$, on the one hand, and of $\mathcal{D}_{\mathrm{TC}}$, on the other hand, happen to resemble each other quite closely, as all of them consists of a sum of functions of the eigenvalues.

The main disadvantage of Huizinga's $\mathcal{D}_{\mathrm{PCA}}$ with respect to $\mathcal{D}_{\mathrm{PCA2}}$ and $\mathcal{D}_{\mathrm{TC}}$ is that it requires to choose a cut-off constant L. For $\mathcal{D}_{\mathrm{PCA2}}$, this user-defined constant is avoided, but the weights j in Equation (4.12) are actually still chosen arbitrarily. For the total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ that we propose, the contribution of each eigenvalue follows naturally from the derivation of mutual information. A key asset of $\mathcal{D}_{\mathrm{TC}}$ is therefore that the influence of each eigenvalue is automatically calibrated, because the expression of the dissimilarity measure is derived from the concept of mutual information.

4.3 Experiments

Huizinga et al. [55] applied the groupwise dissimilarity measures $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$ to a variety of multi-parametric datasets, and compared the results with other state-of-the-art techniques: pairwise mutual information $\mathcal{D}_{\mathrm{MI}}$, the accumulated pairwise estimates (APE) introduced by Wachinger and Navab [150], the groupwise sum of variances designed by Metz et al. [94], and the groupwise mutual information method of Bhatia et al. [14]. Huizinga et al. [55] concluded that their measures $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$ yielded better or equal registration results with respect to the other tested methods.

The present experiment uses total correlation $\mathcal{D}_{\mathrm{TC}}$ as groupwise dissimilarity measure for the registration of the same datasets as in Huizinga et al [55]. On these datasets, the methods of Huizinga et al. [55] were shown to be the best ones, which is why we will compare the registration results of $\mathcal{D}_{\mathrm{TC}}$ with $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$ only.

The results reported by Huizinga et al. [55] for the other dissimilarity measures are directly comparable with the results reported in this chapter.

4.3.1 Description of the six image datasets

The first dataset, denoted CT-LUNG [18], consists of ten patient subsets containing $G = 10$ three-dimensional CT images of the thorax. The intensity distribution in this dynamic imaging dataset are analogous in all images, which means that the model m_g can be considered as a constant in Equation (4.25): it is a particular case of multi-parametric dataset. The second study, denoted T1MOLLI-HEART [139], consists of nine T_1-weighted MRI datasets of porcine hearts with transmural myocardial infarction of the lateral wall. $G = 11$ two-dimensional images were acquired for nine subjects. For each registration case, a voxelwise curve fitting was applied to the registered images, producing quantitative T_1 maps. The third study, denoted T1VFA-CAROTID [24], involves MRIs of the carotid arteries. $G = 5$ three-dimensional images were acquired for 8 human patients. For each patient, the images were registered and fitted to obtain quantitative T_1 maps. The fourth study consists of DW-MR images of the abdominal region, and is denoted ADC-ABDOMEN [48]. Five datasets, each of them including $G = 19$ three-dimensional images, were registered and fitted to produce ADC maps. The fifth study is denoted DTI-BRAIN [30, 77, 111, 141, 152] and consists, for each of the five considered datasets, of registering diffusion tensor images (DTI) of the brain. The number of images to register varied between $G = 33$ and $G = 70$ for each dataset [55]. The fitted parameter is the mean diffusivity (*MD*). The sixth study involves DCE images of the abdomen. Five DCE-ABDOMEN [63] datasets were acquired, each of them containing $G = 160$ three-dimensional images. The fitted parameter of interest considered in this study is K^{trans}. The full descriptions of the fitting models are provided by Huizinga et al. [55].

All human data used in this chapter came from anonymised datasets. Data from the CT-LUNG dataset was obtained from a publicly available dataset[1] [18]. The ethics committee of the Academisch Medisch Centrum, Amsterdam, the Netherlands, approved the research related to the T1VFA-CAROTID and DCE-ABDOMEN datasets. The Research Ethics Committee of the Royal Marsden Hospital, United Kingdom, approved the research related to the ADC-ABDOMEN dataset. The medical ethics committee for research in humans of the University Medical Center Utrecht, the Netherlands, approved the research performed on the DTI-BRAIN dataset. Informed consent was obtained from all patients in human datasets. Porcine data from the T1MOLLI-HEART dataset were approved by the Animal Ethics Committee of the Erasmus MC Rotterdam, the Netherlands. All studies were carried out in accordance with the relevant guidelines and regulations.

4.3.2 Registration characteristics

We selected the same registration settings as Huizinga et al. [55], for comparison purposes. Two resolutions of 1,000 iterations were used for all six image data-

[1] `https://www.dir-lab.com`

sets. To account for deformations caused by heart-pulsations and breathing, we used a B-spline transformation model for the CT-LUNG, T1MOLLI-HEART, T1VFA-CAROTID, ADC-ABDOMEN and DCE-ABDOMEN datasets. The registrations were performed for three distinct B-spline grid spacings: 32 mm, 64 mm and 128 mm for the T1MOLLI-HEART, ADC-ABDOMEN, DCE-ABDOMEN datasets, 8 mm, 16 mm and 32 mm for the T1VFA-CAROTID dataset, and 6 mm, 13 mm and 20 mm for the CT-LUNG dataset. All results are reported as supplementary material (Section 4.7). Results for the intermediate values of the spacings (i.e. either 64 mm, 16 mm or 13 mm), are reported in Section 4.4. To account for deformations caused by head motion and eddy current distortions, we used an affine transformation model for the DTI-BRAIN dataset. When applying $\mathcal{D}_{\mathrm{PCA}}$, the value of L was 1 for CT-LUNG, 3 for T1MOLLI-HEART, 1 for T1VFA-CAROTID, 4 for ADC-ABDOMEN, 7 for DTI-BRAIN, and 4 for DCE-ABDOMEN.

4.3.3 Evaluation measures

No ground truth alignment was available for any of the six datasets considered. However, registration performance was evaluated based on four different measures, described in Huizinga et al. [55], and briefly described in this section.

The first two measures are based on landmark correspondence and overlap of volumes of interest. Landmarks were manually defined on images of the T1VFA-CAROTID and DCE-ABDOMEN datasets. The correspondence between the corresponding landmarks was evaluated by computing a mean target registration error (mTRE). In the T1MOLLI-HEART case, segmentations of the myocardium were outlined on between 6 and 9 images per patient. In the ADC-ABDOMEN case, the spleen was manually delineated on 8 images. For these two cases, the overlap between the segmented structures was then evaluated using a Dice coefficient. For the DTI-BRAIN study, neither landmarks nor structures could be reliably identified on the diffusion weighted images, which is why no overlap or point correspondence was calculated [55].

The second measure quantifies the smoothness of the transformation obtained through registration. Extreme and non-smooth deformations are unexpected. The smoothness of the deformation field can therefore be used to identify such undesirable transformations. A smoothness quantification can be obtained by computing the standard deviation of the determinant of $\partial \boldsymbol{T}_g / \partial \boldsymbol{x}$ over all $\boldsymbol{x}$ for all images: $\mathrm{STD}_{\det(\partial \boldsymbol{T}_g / \partial \boldsymbol{x})}$. Smoothness was quantified for all datasets except for DTI-BRAIN because an affine transformation was used in that last case. The smoother the transformation, the lower the quantity $\mathrm{STD}_{\det(\partial \boldsymbol{T}_g / \partial \boldsymbol{x})}$.

The last evaluation measure is an uncertainty estimation of the qMRI fit. For all datasets, curve fittings were performed to respectively generate T_1, T_1, ADC, MD and K^{trans} quantitative maps. The qMRI models were fitted using a maximum likelihood estimator that takes into account the Rician characteristic of the noise in MR data. We used the fitting same method as Huizinga et al. [55], based on the work of Poot et al. [107]. The uncertainty of these fitted qMRI model parameters can be quantified by the 90$^{\mathrm{th}}$ percentile of the square root of Fréchet-Cramér-Rao lower bound (FCRLB), which provides a lower bound for the variance of the maximum likelihood

parameters. This uncertainty estimate is denoted $90^{\text{th}}\sqrt{\text{FCRLB}}$.

4.3.4 Assessment of multivariate joint normality

As mentioned in Section 4.2, the computation of the total correlation dissimilarity measure $\mathcal{D}_{\text{TC}}$ that we propose is based on the approximation that the intensity distribution of the images to register is multivariate normal. For most datasets, however, the intensity distribution is expected not to be multivariate normal. The underlying idea is that the approximated dissimilarity measure will result in the same minimisation result as if the approximation had not been done.

A second interest of the experimental setting is therefore to evaluate how multivariate normal the intensity distributions are for the six types of datasets that are registered in this chapter, and in the light of the registration accuracy results, to assess whether the approximation that we made can be considered as sensible on multi-parametric datasets.

The joint normality of two images can be assessed by computing and visualizing their joint histogram. Assessing joint normality on more images requires other methods. A possible graphical approach to analyze the multivariate joint normality of G images is to compare the distributions of observed Mahalanobis distances with the distribution of a chi-square distribution with G degrees of freedom χ^2_G. A squared Mahalanobis distance d_i^2 (with $i = 1...N$) can be computed at each voxel location $M_g(i)$, by: $d_i^2 = (\boldsymbol{y}_i - \boldsymbol{y}_m)^T \mathbf{S}^{-1} (\boldsymbol{y}_i - \boldsymbol{y}_m)$, with $\boldsymbol{y}_i = [M_1(i), .., M_G(i)]^T$, the sample mean vector $\boldsymbol{y}_m = \sum_{i=1}^{N} \boldsymbol{y}_i / N$, and the sample covariance $\mathbf{S} = \sum_{i=1}^{N} (\boldsymbol{y}_i - \boldsymbol{y}_m)(\boldsymbol{y}_i - \boldsymbol{y}_m)^T / (N-1)$. It has been shown that the sample squared Mahalanobis distance converges to χ^2_G when $\boldsymbol{y}_i \sim \mathcal{N}_k(\boldsymbol{y}_m, \mathbf{S})$ [137]. To graphically check whether the distribution of intensities of $\mathbf{M}$ is joint normal, we will plot the cumulative distribution function (CDF) of d^2 and χ^2_G in the same graph. If the CDF of the squared Mahalanobis distances d^2 approaches this of χ^2_G, then we will consider the data as joint normal.

4.3.5 Computational efficiency of $\mathcal{D}_{\text{TC}}$

To study the computational efficiency of the proposed total correlation dissimilarity measure $\mathcal{D}_{\text{TC}}$, the average time per iteration is studied by varying three registration parameters: the number of images G that are simultaneously registered, the number of spatial samples taken to evaluate the groupwise dissimilarity measure, and the number of B-spline control points of the transformation model used to warp the images. The influence of these three parameters on the average time per iteration is studied by varying each of them while setting the two remaining ones at values in the range of those described in Section 4.3.2:

- when the number of B-spline control points evolves, the number of images G is set to 50, and the number of spatial samples to 1024. The numbers of B-spline control points per image vary between 50 and 20,000;
- when the number of images G evolves, the number of B-spline control points is set to 500 per image, and the number of spatial samples to 1024. The numbers

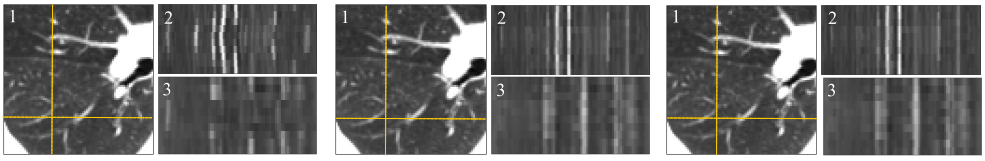
of images G cover the characteristics of the images described in the 'Description of the six image datasets' section (i.e. G = 5...160);
- when the number of spatial samples evolves, the number of B-spline control points is set to 500 per image, and the number of images G is set to 50. We considered numbers of spatial samples between 16 samples and 8,192.

4.4 Results

Groupwise registration based on the total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ that we propose in this study is tested on six different types of image datasets, which overall represents 42 subjects. Dynamic series of CT images were acquired for the first type of image dataset, denoted CT-LUNG. The five other types of datasets, denoted T1MOLLI-HEART, T1VFA-CAROTID, ADC-ABDOMEN, DTI-BRAIN, and DCE-ABDOMEN, are qMRI datasets for which multiple MR images were acquired using different acquisition parameters (or at multiple time points after injection of a contrast agent). For these five qMRI datasets, we fitted a qMRI model to the image intensities at each spatial location, and extracted quantitative images: spin–lattice relaxation time (T_1) images for T1MOLLI-HEART and T1VFA-CAROTID, apparent diffusion coefficient (ADC) images for ADC-ABDOMEN, mean diffusivity (MD) images for DTI-BRAIN, and transfer constant (K^{trans}) images for DCE-ABDOMEN.

4.4.1 Registration accuracy

Figure 4.2 provides a visualisation of the image alignment for one of the CT-LUNG datasets, gathering 10 CT images acquired at different time points from the lung area of a patient.



(a) Misaligned case (b) $\mathcal{D}_{\mathrm{PCA2}}$ (c) $\mathcal{D}_{\mathrm{TC}}$

Figure 4.2: *Registration results for a CT-LUNG dataset. The images denoted '2' and '3' stack the voxel information of G = 10 images at the locations defined by the yellow dotted lines drawn in the image denoted '1' (vertical line: '2', horizontal line: '3').*

Misalignments due to respiratory motion are visible when no registration is applied between the images (Figure 4.2a), while they disappear after applying image registration based on Huizinga's $\mathcal{D}_{\mathrm{PCA2}}$ (Figure 4.2b) or on the total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ proposed in this chapter (Figure 4.2c). Visual differences between the results obtained with $\mathcal{D}_{\mathrm{PCA2}}$ and $\mathcal{D}_{\mathrm{TC}}$ are more limited and harder to identify.

For the five qMRI datasets, Figure 4.3 provides quantitative parameter images obtained by applying curve fitting to the images before registration, after registration using Huizinga's $\mathcal{D}_{\mathrm{PCA2}}$ groupwise dissimilarity measure, and after registration using the total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ proposed in this chapter. The fitting models used to derive the qMRI images assume that spatial correspondence is ensured between the images used for curve fitting. It is therefore expected that quantitative images obtained after image registration will be more reliable than without image registration [48, 55]. Based on Figure 4.3, visual differences in the estimates tissue maps are easily noticeable between the case before image registration, on the one hand, and the cases with $\mathcal{D}_{\mathrm{PCA2}}$ or $\mathcal{D}_{\mathrm{TC}}$, on the other hand. Such differences are particularly visible at organ interfaces. Slighter changes, identified by green arrows, can be identified between the tissue maps obtained with $\mathcal{D}_{\mathrm{PCA2}}$ and $\mathcal{D}_{\mathrm{TC}}$.

Full registration accuracy results in terms of landmark/volume correspondence (mTRE or Dice coefficient), registration transformation smoothness $\mathrm{STD}_{\det(\partial \boldsymbol{T}_g/\partial \mathbf{x})}$, and uncertainty estimation FCRLB, are provided as supplementary material (Section 4.7) for the following dissimilarity measures: pairwise mutual information $\mathcal{D}_{\mathrm{MI}}$, Huizinga's dissimilarity measures based on PCA $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$, and the total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$.

Table 4.1 presents a partial version of the registration accuracy results, based on the middle value of the control point spacings that were used for the non-rigid B-spline transformation model: 13 mm for CT-LUNG, 64 mm for T1MOLLI-HEART, 16 mm for T1VFA-CAROTID, 64 mm for ADC-ABDOMEN, and 64 mm for DCE-ABDOMEN. Registration performances in terms of landmark correspondence (mean target registration error, denoted mTRE) or overlap of volumes of interest (Dice coefficients) are given in Table 4.1a. For all dataset, better alignments (i.e. lower mTRE) or overlaps (i.e. higher Dice coefficients) were obtained with the groupwise measures $\mathcal{D}_{\mathrm{TC}}$, $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$ than with pairwise mutual information $\mathcal{D}_{\mathrm{MI}}$, with one exception: the mTRE obtained with $\mathcal{D}_{\mathrm{PCA2}}$ for the CT-LUNG dataset is higher than the mTRE obtained with $\mathcal{D}_{\mathrm{MI}}$. The Dice coefficients and mTRE results are very similar for $\mathcal{D}_{\mathrm{TC}}$, $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$. The only case for which $\mathcal{D}_{\mathrm{TC}}$ performs slightly worse than the two other groupwise measures is on the DCE-ABDOMEN dataset. Table 4.1b provides values for the transformation smoothness $\mathrm{STD}_{\det(\partial \boldsymbol{T}_g/\partial \mathbf{x})}$. In all cases, $\mathcal{D}_{\mathrm{TC}}$, $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$ yield lower (i.e. better) values of $\mathrm{STD}_{\det(\partial \boldsymbol{T}_g/\partial \mathbf{x})}$ than $\mathcal{D}_{\mathrm{MI}}$. The only case for which $\mathcal{D}_{\mathrm{TC}}$ performs slightly worse than the two other groupwise measures is on the T1VFA-CAROTID dataset. Table 4.1c provides uncertainty estimations of the qMRI fit ($90^{\mathrm{th}}\sqrt{\mathrm{FCRLB}}$). The results indicate that the values of $90^{\mathrm{th}}\sqrt{\mathrm{FCRLB}}$ are lower (i.e. better) with $\mathcal{D}_{\mathrm{TC}}$ than with $\mathcal{D}_{\mathrm{MI}}$ for the T1MOLLI-HEART and DCE-ABDOMEN datasets, while they are quite similar for T1VFA-CAROTID and DTI-BRAIN, and higher (i.e. worse) for the ADC-ABDOMEN dataset. The $90^{\mathrm{th}}\sqrt{\mathrm{FCRLB}}$ obtained with $\mathcal{D}_{\mathrm{TC}}$ is higher than the $90^{\mathrm{th}}\sqrt{\mathrm{FCRLB}}$ obtained with $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$ for two datasets (ADC-ABDOMEN and DCE-ABDOMEN), while it is similar or better for three datasets (T1MOLLI-HEART, T1VFA-CAROTID, and DTI-BRAIN). The full results (Tables S1 to S6) are consistent with the results presented in Tables 4.1a-c.

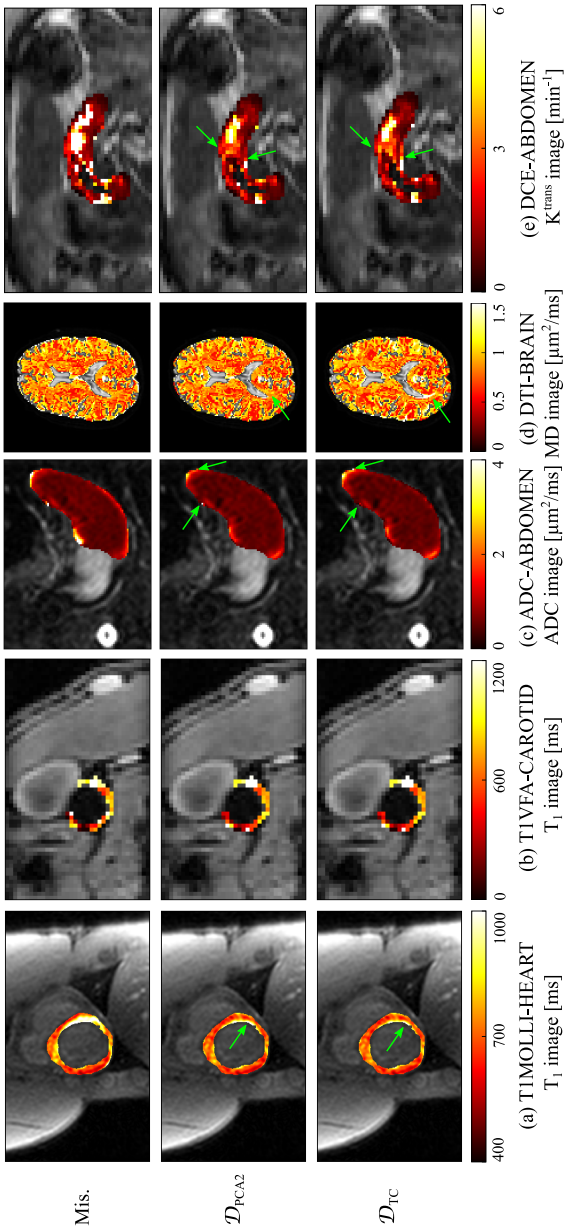


Figure 4.3: *Tissue maps generated before image registration (top), after image registration with $\mathcal{D}_{\mathrm{PCA2}}$ (middle), and after image registration with $\mathcal{D}_{\mathrm{TC}}$ (bottom). The fitted values are shown in the myocardium for T1MOLLI-HEART, in the carotid artery wall for T1VFA-ABDOMEN, in the spleen for ADC-ABDOMEN, in the brain parenchyma for DTI-BRAIN, and in the pancreas for DCE-ABDOMEN. Slight visual changes between the tissue maps obtained with $\mathcal{D}_{\mathrm{PCA2}}$ and $\mathcal{D}_{\mathrm{TC}}$ are identified by green arrows.*

Table 4.1: *Registration accuracy, smoothness and uncertainty results obtained with the middle value of the control point spacings for the non-rigid B-spline transformation model. The full results are presented in Section 4.7.*

Dataset	CT LUNG	T1MOLLI HEART	T1VFA CAROTID	ADC ABDOMEN	DTI BRAIN	DCE ABDOMEN
Measure	*mTRE* [mm]	*Dice* [%]	*mTRE* [mm]	*Dice* [%]	- -	*mTRE* [mm]
Mis.	6.72 ± 2.51	48 ± 8	1.47 ± 0.54	70 ± 4	-	8.49 ± 4.54
$\mathcal{D}_{\mathrm{MI}}$	1.43 ± 0.23	37 ± 11	1.22 ± 0.43	64 ± 16	-	6.46 ± 2.32
$\mathcal{D}_{\mathrm{PCA}}$	1.40 ± 0.37	53 ± 7	1.11 ± 0.42	71 ± 5	-	6.11 ± 2.33
$\mathcal{D}_{\mathrm{PCA2}}$	1.56 ± 0.55	52 ± 11	1.08 ± 0.39	75 ± 5	-	5.99 ± 2.18
$\mathcal{D}_{\mathrm{TC}}$	1.42 ± 0.40	53 ± 11	1.09 ± 0.40	74 ± 5	-	6.18 ± 2.40

(a) Dice coefficients or mTRE values
(mean value ± standard deviation)

Dataset	CT LUNG	T1MOLLI HEART	T1VFA CAROTID	ADC ABDOMEN	DTI BRAIN	DCE ABDOMEN
Mis.	0 ± 0	0 ± 0	0 ± 0	0 ± 0	-	0 ± 0
$\mathcal{D}_{\mathrm{MI}}$	15 ± 4	7 ± 2	2 ± 0	8 ± 3	-	4 ± 2
$\mathcal{D}_{\mathrm{PCA}}$	8 ± 2	2 ± 1	2 ± 1	3 ± 2	-	4 ± 2
$\mathcal{D}_{\mathrm{PCA2}}$	7 ± 2	1 ± 1	1 ± 0	3 ± 1	-	2 ± 1
$\mathcal{D}_{\mathrm{TC}}$	8 ± 2	2 ± 1	1 ± 0	5 ± 2	-	4 ± 2

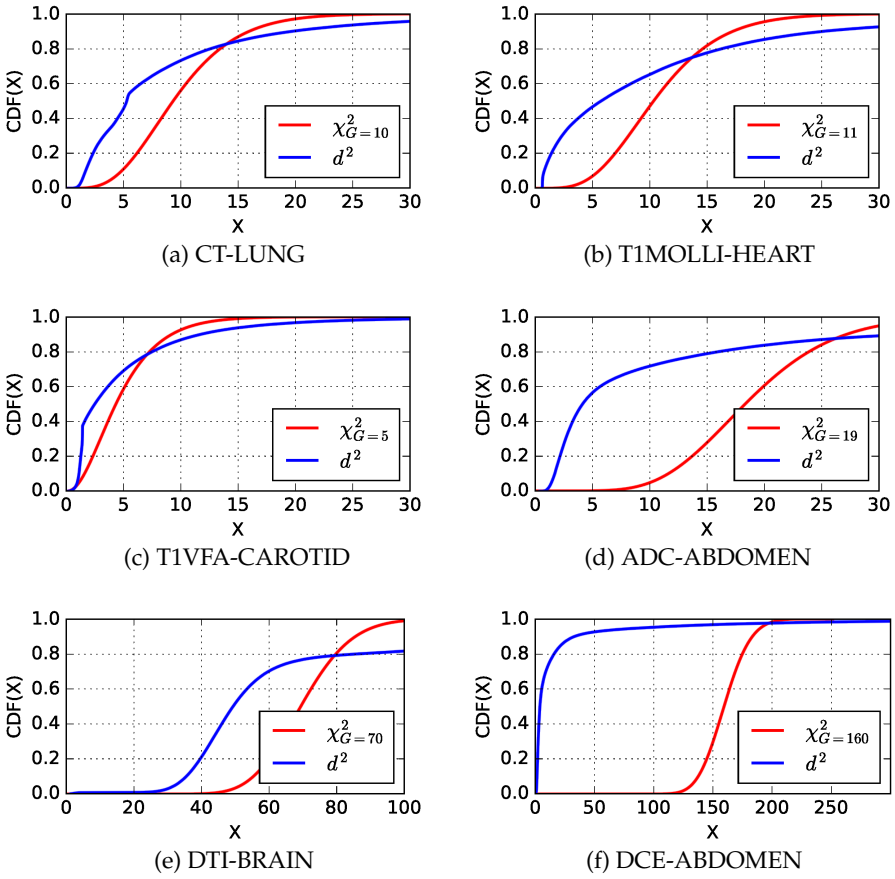
(b) Transformation smoothness $\mathrm{STD}_{\det(\partial \boldsymbol{T}_g / \partial \boldsymbol{x})}$ [%]
(mean value ± standard deviation)

Dataset	CT LUNG	T1MOLLI HEART	T1VFA CAROTID	ADC ABDOMEN	DTI BRAIN	DCE ABDOMEN
Parameter	- -	T_1 [ms]	T_1 [ms]	*ADC* [μm^2/ms]	*MD* [μm^2/ms]	K^{trans} [min^{-1}]
Mis.	-	92 ± 19	> 1000	1.37 ± 0.83	0.096 ± 0.029	2.84 ± 2.30
$\mathcal{D}_{\mathrm{MI}}$	-	97 ± 16	501 ± 83	0.25 ± 0.05	0.084 ± 0.028	3.64 ± 4.13
$\mathcal{D}_{\mathrm{PCA}}$	-	87 ± 16	498 ± 93	0.23 ± 0.06	0.085 ± 0.029	1.52 ± 1.18
$\mathcal{D}_{\mathrm{PCA2}}$	-	83 ± 12	510 ± 110	0.27 ± 0.05	0.084 ± 0.028	1.27 ± 0.92
$\mathcal{D}_{\mathrm{TC}}$	-	77 ± 13	500 ± 96	0.32 ± 0.05	0.085 ± 0.028	1.87 ± 1.79

(c) Uncertainty estimation $90^{\mathrm{th}} \sqrt{\mathrm{FCRLB}}$ of the fitted parameters
(mean value ± standard deviation)

4.4.2 Multivariate joint normality

As detailed in Section 4.2, the computation of the total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ that we propose is based on the approximation that the intensity distribution of the images to register is multivariate normal. Cumulative distribution functions (CDF) of the squared Mahalanobis distance d^2, representing the intensity distribution for each of the six dataset types, are plotted in Figure 4.4. According to these plots, none of these measure CDF follows the theoretical multivariate normal CDF (χ^2_G distribution), which suggests that the image intensities of the images do not follow a multivariate normal distribution.



(a) CT-LUNG (b) T1MOLLI-HEART

(c) T1VFA-CAROTID (d) ADC-ABDOMEN

(e) DTI-BRAIN (f) DCE-ABDOMEN

Figure 4.4: *Cumulative distribution functions for one subject of the six image dataset types (aligned case). The observed CDF (blue) is compared with the theoretical CDF of a chi-square distribution with G degrees of freedom (red).*

4.4.3 Computational efficiency of $\mathcal{D}_{\mathrm{TC}}$

Figure 4.5 illustrates the evolution of the average time per iteration obtained with groupwise total correlation $\mathcal{D}_{\mathrm{TC}}$ for three image registration parameters: the number of B-spline control points per image, the number of images G, and the number of spatial samples taken to evaluate the dissimilarity measure. The results show that the average registration time per iteration monotonically increases with each of the considered registration parameter. With the present implementation of $\mathcal{D}_{\mathrm{TC}}$ and of the registration components of the elastix software used to perform the registrations, the results indicate that the number of B-spline control points has a limited influence on the average time per iteration as it remains close to 9 seconds for the whole span of numbers of B-spline control points that we considered. The experiments suggest that the number of images G influences the computation time most. For instance,

when the number of image is $G = 40$, the average iteration time is 5 seconds, while this time reaches about two minutes for $G = 160$ images. In terms of the number of spatial samples, multiplying the number of spatial samples by 4 ends up in an average time per iteration that is multiplied by 6.

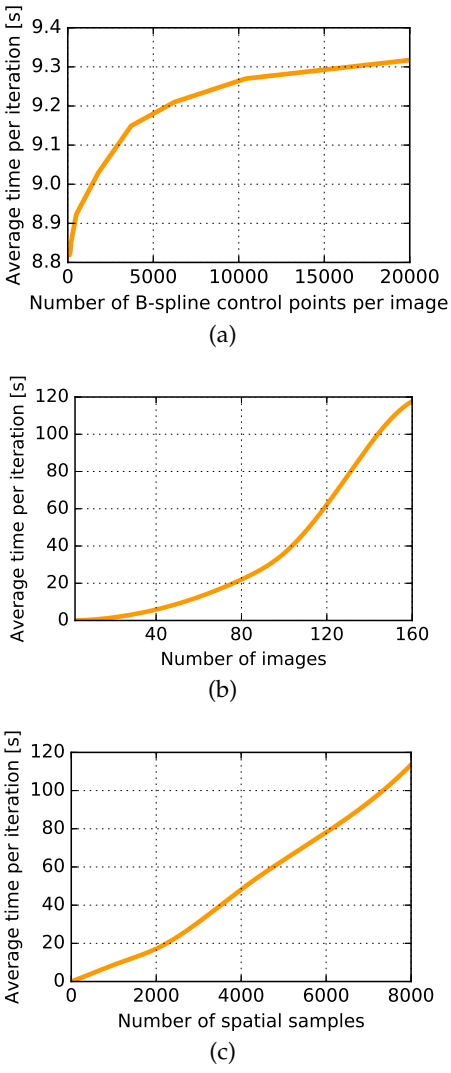


Figure 4.5: *Average time per iteration with respect to the number of B-spline control points per image (a), the number of images G (b), and the number of spatial samples (c).*

4.5 Discussion

The focus of this chapter was to adapt a multivariate version of mutual information in the context of the groupwise registration of medical images, so that it can be used to register two or more images in one optimisation procedure.

Among the main multivariate versions of mutual information, namely interaction information $\mathcal{D}_{\mathrm{II}}$, total correlation $\mathcal{D}_{\mathrm{TC}}$ and dual total correlation $\mathcal{D}_{\mathrm{DTC}}$, total correlation $\mathcal{D}_{\mathrm{TC}}$ theoretically allows to quantify the shared information between any subset of the images to register. Besides, the expression of total correlation is particularly straightforward to apply for the registration of $G \geq 2$ images, provided that the image intensity distribution is approximated by a multivariate normal distribution.

The expression of the approximated total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ that we devise is remarkably analogous to the expressions of two other dissimilarity measures $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$ introduced by Huizinga et al. [55], which were developed based on the intuition that an aligned set of images can be described by a small number of high eigenvalues. The expressions of these dissimilarity measures are all sums of functions of the eigenvalues of the correlation matrix $\mathbf{K}$ (compare Equations (4.18), (4.26) and (4.27)). Huizinga et al. [55] had proposed to weigh more the last eigenvalues (the λ_i with the highest i indexes) than the first ones (the λ_i with the lowest i indexes) so that as much variance as possible is explained by a few large eigenvectors. The form of $\mathcal{D}_{\mathrm{TC}}$ obtained in this chapter confirms the intuition of Huizinga et al. [55], since the natural logarithm in Equation (4.18) also puts more weight on the lower eigenvalues than on the higher ones.

Results obtained on multi-parametric datasets show that the total correlation method that we propose yields comparable results as the PCA-based methods of Huizinga et al. [55], and better registration results than pairwise mutual information $\mathcal{D}_{\mathrm{MI}}$. The main advantage of $\mathcal{D}_{\mathrm{TC}}$ with respect to $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$ is that it is more theoretically justified: the contribution of each eigenvalue used to compute $\mathcal{D}_{\mathrm{TC}}$ is automatically calibrated and follows naturally from the concepts of multivariate mutual information. For the two other dissimilarity measures, the eigenvalue calibration are set empirically.

This chapter shows that even though the intensity distribution of the datasets to register is not multivariate normal, $\mathcal{D}_{\mathrm{TC}}$ yields registration results that are better than mutual information and similar to the PCA dissimilarity measures of Huizinga et al. [55]. This is the case for a total of six diverse multi-parametric datasets, which suggests that approximating the intensity distributions, as done in this chapter, yields optimisation minima that result in comparable or better registration accuracies than other state-of-the-art pairwise and groupwise techniques. On multi-parametric datasets, the results suggest that the approximation by a multivariate normal distribution is not detrimental to the registration results.

In the current implementation of the total correlation registration technique, increases in the number of images G have the largest impact on the average time per iteration, which is not surprising as both the amount of image data to register and the number of transformations to estimate scale with a factor G; moreover, estimating the correlation matrix $\mathbf{K}$ and its eigenvalue decomposition become increasingly computationally demanding. Further optimisations could improve the scalability of

total correlation with respect to the number of images. The computation time also scales linearly with the number of spatial samples. Thanks to the use of the stochastic gradient descent optimisation routine, we were able to use a relatively low number (2048) of spatial samples in our experiments, while still achieving accurate registration.

4.6 Conclusion

We proposed an implementation of an approximated version of total correlation $\mathcal{D}_{\mathrm{TC}}$ for which the registration results are comparable to the results obtained with the dissimilarity measures of Huizinga et al. [55], on multi-parametric datasets. Our results indicate that approximating the intensity distributions by a joint normal distribution for the sake of efficient calculation of the entropy, used to derive total correlation $\mathcal{D}_{\mathrm{TC}}$, does not constitute a limitation in the practical application of $\mathcal{D}_{\mathrm{TC}}$ to multi-parametric image datasets. Total correlation $\mathcal{D}_{\mathrm{TC}}$ has the advantage of being elegant and theoretically justified, while the dissimilarity measures $\mathcal{D}_{\mathrm{PCA}}$ and $\mathcal{D}_{\mathrm{PCA2}}$ proposed by Huizinga et al. [55] were elaborated empirically. Additionally, groupwise total correlation $\mathcal{D}_{\mathrm{TC}}$ offers an alternative to pairwise registration based on mutual information on multi-parametric imaging datasets.

4.7 Supplementary material

The tables reported in this section report the full results for the datasets CT-LUNG (Table 4.2), T1MOLLI-HEART (Table 4.3), T1VFA-CAROTID (Table 4.4), ADC-ABDOMEN (Table 4.5), DTI-BRAIN (Table 4.6) and DCE-ABDOMEN (Table 4.7). Columns with a gray background were already reported in Table 4.1.

Table 4.2: *CT-LUNG. Mean and standard deviation over all subjects.*

Point spacing	6 mm	13 mm	20 mm
Mis.	6.72 ± 2.51	6.72 ± 2.51	6.72 ± 2.51
$\mathcal{D}_{\text{MI}}$	1.78 ± 0.40	1.43 ± 0.23	1.45 ± 0.21
$\mathcal{D}_{\text{PCA}}$	1.47 ± 0.60	1.40 ± 0.37	1.45 ± 0.34
$\mathcal{D}_{\text{PCA2}}$	1.72 ± 0.78	1.56 ± 0.55	1.59 ± 0.49
$\mathcal{D}_{\text{TC}}$	1.48 ± 0.62	1.42 ± 0.40	1.47 ± 0.37

(a) mTRE [mm]

Point spacing	6 mm	13 mm	20 mm
Mis.	0 ± 0	0 ± 0	0 ± 0
$\mathcal{D}_{\text{MI}}$	28 ± 8	15 ± 4	11 ± 3
$\mathcal{D}_{\text{PCA}}$	12 ± 3	8 ± 2	7 ± 2
$\mathcal{D}_{\text{PCA2}}$	10 ± 2	7 ± 2	6 ± 1
$\mathcal{D}_{\text{TC}}$	12 ± 3	8 ± 2	7 ± 2

(b) $\text{STD}_{|\partial \mathbf{T} g / \partial \boldsymbol{x}|}$ [%]

Table 4.3: *T1MOLLI-HEART. Mean and standard deviation over all subjects for reference images 1, 4, 7 and 11.*

Point spacing	32 mm				64 mm				128 mm			
Ref. image	1	4	7	11	1	4	7	11	1	4	7	11
Mis.	48± 8	48± 8	48± 8	48± 8	48± 8	48± 8	48± 8	48± 8	48± 8	48± 8	48± 8	48± 8
$\mathcal{D}_{\text{MI}}$	26±11	22±10	22±15	39±14	37±11	34±16	19±15	52± 7	44±14	40±13	24±17	51± 7
$\mathcal{D}_{\text{PCA}}$	54± 9	54± 9	54± 9	54± 9	53± 7	53± 7	53± 7	53± 8	51± 9	50± 9	51± 9	51± 9
$\mathcal{D}_{\text{PCA2}}$	52±13	52±13	52±13	52±13	52±11	52±10	52±11	53±10	52± 7	52± 8	52± 8	52± 8
$\mathcal{D}_{\text{TC}}$	52±13	52±12	52±13	53±13	53±11	53±11	53±11	53±11	53± 7	53± 8	54± 7	53± 8

(a) Dice coefficient [%]

Point spacing	32 mm				64 mm				128 mm			
Ref. image	1	4	7	11	1	4	7	11	1	4	7	11
Mis.	0± 0	0± 0	0± 0	0± 0	0± 0	0± 0	0± 0	0± 0	0± 0	0± 0	0± 0	0± 0
$\mathcal{D}_{\text{MI}}$	24± 8	28± 8	29± 9	10± 4	7± 2	7± 2	8± 3	3± 1	1± 0	1± 0	1± 0	1± 0
$\mathcal{D}_{\text{PCA}}$	6± 4	4± 2	4± 2	4± 1	2± 1	1± 1	2± 1	1± 1	0± 0	0± 0	0± 0	0± 0
$\mathcal{D}_{\text{PCA2}}$	6± 2	5± 2	6± 2	4± 1	1± 1	1± 1	2± 1	1± 1	0± 0	0± 0	0± 0	0± 0
$\mathcal{D}_{\text{TC}}$	7± 3	6± 3	7± 3	5± 1	2± 1	2± 1	2± 1	1± 1	0± 0	0± 0	0± 0	0± 0

(b) $\text{STD}_{|\partial \mathbf{T} g / \partial \boldsymbol{x}|}$ [%]

Point spacing	32 mm				64 mm				128 mm			
Ref. image	1	4	7	11	1	4	7	11	1	4	7	11
Mis.	92±19	92±19	92± 19	92±19	92±19	92±19	92±19	92±19	92±19	92±19	92±19	92±19
$\mathcal{D}_{\text{MI}}$	119±26	157±61	467±610	93±19	97±16	103±18	146±52	83±13	91±14	95±16	121±29	81±12
$\mathcal{D}_{\text{PCA}}$	83±17	79±17	78± 11	85±16	87±16	84±20	83±13	85±13	88±14	89±21	103±50	85±10
$\mathcal{D}_{\text{PCA2}}$	81±15	79±16	79± 12	79±12	83±12	79±14	80±10	81±10	84±12	82±15	82±10	81±11
$\mathcal{D}_{\text{TC}}$	77±14	75±11	73± 14	74±12	77±13	75±11	73±13	74±11	79±12	76±11	76±13	75±11

(c) $90\% \sqrt{\text{FCRLB}}\, T_1$ [ms]

Table 4.4: *T1VFA-CAROTID. Mean and standard deviation.*

Point spacing	8 mm	16 mm	32 mm
Mis.	1.47±0.54	1.47±0.54	1.47±0.54
$\mathcal{D}_{\mathrm{MI}}$	1.26±0.44	1.22±0.43	1.23±0.45
$\mathcal{D}_{\mathrm{PCA}}$	1.25±0.56	1.11±0.42	1.10±0.43
$\mathcal{D}_{\mathrm{PCA2}}$	1.13±0.46	1.08±0.39	1.10±0.43
$\mathcal{D}_{\mathrm{TC}}$	1.19±0.50	1.09±0.40	1.10±0.42

(a) mTRE [mm]

Point spacing	8 mm	16 mm	32 mm
Mis.	0±0	0±0	0±0
$\mathcal{D}_{\mathrm{MI}}$	7±1	2±0	0±0
$\mathcal{D}_{\mathrm{PCA}}$	6±1	2±1	0±0
$\mathcal{D}_{\mathrm{PCA2}}$	5±1	1±0	0±0
$\mathcal{D}_{\mathrm{TC}}$	6±1	1±0	0±0

(b) $\mathrm{STD}_{|\partial \mathbf{T} g / \partial \boldsymbol{x}|}$ [%]

Point spacing	8 mm	16 mm	32 mm
Mis.	> 1000	> 1000	> 1000
$\mathcal{D}_{\mathrm{MI}}$	530±136	501± 83	523± 93
$\mathcal{D}_{\mathrm{PCA}}$	540±154	498± 93	530± 94
$\mathcal{D}_{\mathrm{PCA2}}$	532±154	510±110	523± 87
$\mathcal{D}_{\mathrm{TC}}$	533±150	500± 96	528± 96

(c) 90%$\sqrt{\mathrm{FCRLB}}$ T_1 [ms]

Table 4.5: *ADC-ABDOMEN. Mean and standard deviation.*

Point spacing	8 mm	16 mm	32 mm
Mis.	70± 4	70± 4	70±4
$\mathcal{D}_{\mathrm{MI}}$	61±18	64±16	73±8
$\mathcal{D}_{\mathrm{PCA}}$	65±13	71± 5	71±4
$\mathcal{D}_{\mathrm{PCA2}}$	75± 7	75± 5	73±5
$\mathcal{D}_{\mathrm{TC}}$	72± 6	74± 5	73±5

(a) mTRE [mm]

Point spacing	8 mm	16 mm	32 mm
Mis.	0± 0	0±0	0±0
$\mathcal{D}_{\mathrm{MI}}$	25±11	8±3	2±1
$\mathcal{D}_{\mathrm{PCA}}$	11± 3	3±1	1±0
$\mathcal{D}_{\mathrm{PCA2}}$	8± 2	3±1	0±0
$\mathcal{D}_{\mathrm{TC}}$	16± 5	5±2	1±0

(b) $\mathrm{STD}_{|\partial \mathbf{T} g / \partial \boldsymbol{x}|}$ [%]

Point spacing	8 mm	16 mm	32 mm
Mis.	0.24±0.06	0.25±0.05	0.29±0.05
$\mathcal{D}_{\mathrm{MI}}$	0.20±0.08	0.23±0.10	0.35±0.16
$\mathcal{D}_{\mathrm{PCA}}$	0.15±0.04	0.23±0.06	0.46±0.27
$\mathcal{D}_{\mathrm{PCA2}}$	0.23±0.03	0.27±0.05	0.41±0.18
$\mathcal{D}_{\mathrm{TC}}$	0.33±0.05	0.32±0.05	0.38±0.14

(c) 90%$\sqrt{\mathrm{FCRLB}}$ ADC [μm^2/ms]

Table 4.6: *DTI-BRAIN. Mean and standard deviation of the 90%$\sqrt{\mathrm{FCRLB}}$ MD [μm^2/ms]*

	Affine
Mis.	0.096±0.029
$\mathcal{D}_{\mathrm{MI}}$	0.084±0.028
$\mathcal{D}_{\mathrm{PCA}}$	0.085±0.029
$\mathcal{D}_{\mathrm{PCA2}}$	0.084±0.028
$\mathcal{D}_{\mathrm{TC}}$	0.085±0.029

Table 4.7: *DCE-ABDOMEN. Mean and standard deviation.*

Point spacing	8 mm	16 mm	32 mm
Mis.	8.49±4.54	8.49±4.54	8.49±4.54
$\mathcal{D}_{\mathrm{MI}}$	6.73±2.02	6.46±2.32	6.47±2.37
$\mathcal{D}_{\mathrm{PCA}}$	6.21±2.25	6.11±2.32	6.24±2.37
$\mathcal{D}_{\mathrm{PCA2}}$	5.89±2.23	5.99±2.17	6.18±2.27
$\mathcal{D}_{\mathrm{TC}}$	7.01±1.82	6.18±2.40	6.23±2.30

(a) mTRE [mm]

Point spacing	8 mm	16 mm	32 mm
Mis.	0±0	0±0	0±0
$\mathcal{D}_{\mathrm{MI}}$	20±9	4±2	1±1
$\mathcal{D}_{\mathrm{PCA}}$	11±4	4±2	1±0
$\mathcal{D}_{\mathrm{PCA2}}$	6±3	2±1	0±0
$\mathcal{D}_{\mathrm{TC}}$	11±4	4±2	1±0

(b) $\mathrm{STD}_{|\partial \mathbf{T} g / \partial \boldsymbol{x}|}$ [%]

Point spacing	8 mm	16 mm	32 mm
Mis.	2.84±2.30	2.84±2.30	2.84±2.30
$\mathcal{D}_{\mathrm{MI}}$	3.85±2.41	3.64±4.13	2.54±2.58
$\mathcal{D}_{\mathrm{PCA}}$	1.69±1.48	1.52±1.18	1.46±1.09
$\mathcal{D}_{\mathrm{PCA2}}$	1.17±0.87	1.27±0.92	1.38±1.16
$\mathcal{D}_{\mathrm{TC}}$	1.78±1.74	1.87±1.81	1.54±1.38

(c) 90%$\sqrt{\mathrm{FCRLB}}$ K^{trans} [min^{-1}]

Groupwise multi-channel image registration

Abstract — Multi-channel image registration is an important challenge in medical image analysis. Multi-channel images may result from modalities that generate images composed of multiple channels, such as dual-energy CT or multispectral microscopy. Multi-channel images can also be built using images from different acquisitions (e.g. different modalities or time points) or through post-acquisition operations (e.g. filtering, computation of feature images). Multi-channel registration techniques have been proposed, but most of them are applicable to only two multi-channel images at a time. In the present chapter, we propose to formulate multi-channel registration as a groupwise image registration problem. In this way, we derive a method that allows the registration of two or more multi-channel images in a fully symmetric manner (i.e. all images play the same role in the registration procedure), and therefore has transitive consistency by definition. The method that we introduce is applicable to any number of multi-channel images, any number of channels per image, and it allows to take into account correlation between any pair of images and not just corresponding channels. In addition, it is fully modular in terms of dissimilarity measure, transformation model, regularisation method, and optimisation strategy. For two multimodal datasets, we computed feature images from the initially acquired images, and applied the proposed registration technique to the newly created sets of multi-channel images. MIND descriptors were used as feature images, and we chose total correlation as groupwise dissimilarity measure. Results show that groupwise multi-channel image registration is a competitive alternative to the pairwise multi-channel scheme, in terms of registration accuracy and insensitivity towards registration reference spaces.

Based upon: J.-M. Guyader, W. Huizinga, V. Fortunati, D. H. J. Poot, J. F. Veenland, M. M. Paulides, W. J. Niessen and S. Klein, "Groupwise multi-channel image registration", *IEEE Journal of Biomedical and Health Informatics*, *in press*.

5.1 Introduction

Image registration is an important tool for medical image analysis. Medical image datasets can be made up of images obtained from different modalities, time points or patients, for instance. Spatial correspondence between the images is therefore, in most cases, not naturally ensured. Automated image registration can be used to realign such datasets, which is crucial for many post-acquisition image processing techniques.

Among the large range of image registration techniques that have been developed, various research works have focused on multi-channel registration. Multi-channel image registration consists of applying registration to images for which several channels are available. The channels of a given multi-channel image can be obtained due to the modality itself (e.g. dual-energy CT or multispectral microscopy), from different post-acquisition operations (e.g. filtering, computation of feature images), or from different acquisitions (e.g. different modalities or time points). In this study, we consider an image as being multi-channel when its channel images are spatially registered, either because they are naturally spatially registered, or thanks to a preliminary registration step. Multi-channel image registration is the task of finding spatial correspondences between several such multi-channel images. Rhode et al. [113, 114] were among the first to propose a method handling the registration of two multi-channel images. They designed an approach based on multivariate correlation, and applied it in the context of diffusion-tensor imaging. Other applications of multi-channel registration focus on the registration of feature images: instead of applying registration to the images originally present in a dataset, feature images derived from the original images are used for the registration. Such a technique was proposed by Legg et al. [78], who extracted several feature images based on Gaussian derivatives, and subsequently registered these feature images using a dissimilarity measure based on regional mutual information [119]. Staring et al. [128] designed α-mutual information, a technique that registers sets of feature images using the concepts of mutual information. Heinrich et al. [53] created a registration similarity measure based on a sum of squared absolute differences of the feature images, to register sets of multiple feature images obtained with a descriptor called modality independent neighbourhood descriptor (MIND). Li et al. [80] derived another descriptor from MIND, and named it the autocorrelation of local structure (ALOST). Suarez et al. [129] used a registration technique based on a dissimilarity measure derived from the correlation matrix of two multi-channel images. Heinrich et al. [54] introduced local canonical correlation analysis, a method that assess dissimilarity into new bases that best represent the relations between two multi-channel images. Chen et al. [20] addressed the pairwise registration of two images of different modalities by generating synthetic images that are considered as a second channel for pairwise image registration.

Most multi-channel methods that were previously proposed are pairwise registration schemes. This means that they are applicable to only two multi-channel images at a time, and require to select a fixed reference image to which the remaining multi-channel image will be registered. These multi-channel pairwise registration schemes have the drawback that they require the choice of fixed reference image,

which may bias registration accuracy, as shown by Geng et al. [45].

In the present study, we propose a novel framework for multi-channel image registration. The key idea is to cast multi-channel registration as a groupwise registration problem. The novel registration technique that we devise is suitable not only for the common case of two multi-channel images, but also for cases in which the aim is to register three or more multi-channel images at a time. The multi-channel registration method that we propose is designed as a groupwise registration problem: it is symmetric and all image data is taken into account simultaneously in a single optimisation procedure.

Various groupwise registration methods have been proposed. One of the earliest groupwise method was the technique of Joshi et al. [60], based on unbiased diffeomorphic registration. Learned-Miller [76] presented a groupwise method consisting of minimising the sum of the pixelwise entropies computed at each voxel location. Bhatia et al. [14] designed a groupwise dissimilarity measure based on the sum of the entropies of each image. Metz et al. [94] proposed a groupwise dissimilarity measure based on voxelwise variance, designed for monomodal images. A few groupwise methods for multimodal images or multiparametric images (i.e. same modality, different acquisition settings) were presented, such as an efficient joint entropy minimisation presented by Spiclin et al. [126], a sum of accumulated pairwise estimates presented by Wachinger et al. [150], a technique based on principal component analysis introduced by [55], a technique based on Pythagorean means created by Polfliet et al. [106], and a technique based on total correlation that we previously proposed [49]. It was not investigated in the literature whether and how these methods could be applicable to multi-channel images.

In this study, the generic groupwise multi-channel registration framework that we propose is applicable to any number of multi-channel images, any number of channels per image, and allows to take into account correlations between any pair of channel images. Moreover, it is fully modular in terms of dissimilarity measure, transformation model, regularisation method and optimisation strategy.

5.2 Materials and methods

5.2.1 Preliminaries

Let us consider $\tilde{M}_1, ..., \tilde{M}_G$, a series of $G \geq 2$ multi-channel images. One multi-channel image will be denoted $\tilde{M}_g$, and its associated channel images will be denoted $\tilde{M}_{g,f}$ (the index f varies between 1 and the number of channels F_g of image $\tilde{M}_g$). The multi-channel images $\tilde{M}_1, ..., \tilde{M}_G$ have $F_1, ..., F_G$ channels, respectively. The complete set of image channels to register is therefore:

$$\tilde{\mathcal{M}} = \left\{ \tilde{M}_{g,f} \text{ with } g = 1...G \text{ and } f = 1...F_g \right\} \tag{5.1}$$

Each channel image $\tilde{M}_{g,f}$ is associated with a differentiable function $\tilde{M}_{g,f}(\boldsymbol{x})$ of the spatial coordinate $\boldsymbol{x}$.

Multi-channel image registration is the task of finding the spatial correspondence between several misaligned multi-channel images, assuming that the channel images of each multi-channel image are already all aligned. This study focuses on intensity-based registration methods based on the minimisation of a dissimilarity measure $\mathcal{D}$.

For the particular case of mono-channel image registration (i.e. $F_1 = ... = F_G = 1$), each image $\tilde{M}_g$ can be assimilated with its single channel image $\tilde{M}_{g,1}$. In that case, the following shorthand notation will be used: $M_g = \tilde{M}_g = \tilde{M}_{g,1}$.

5.2.2 Existing pairwise multi-channel image registration

All multi-channel registration methods designed so far are pairwise: they are limited to $G = 2$ images and require the selection of a fixed reference space among the multi-channel images to register. Let us consider two multi-channel images $\tilde{M}_1$ and $\tilde{M}_2$. When the numbers of channels F_1 and F_2 are equal ($F_1 = F_2 = F$), a generalised pairwise multi-channel (PMC) dissimilarity measure $\mathcal{D}_{\mathrm{PMC}}$ can be written:

$$\mathcal{D}_{\mathrm{PMC}}(\tilde{\mathcal{M}}) = \sum_{f=1}^{F} \mathcal{D}_{\mathrm{pairwise}} \left(\tilde{M}_{1,f}, \tilde{M}_{2,f} \right) \tag{5.2}$$

with $\mathcal{D}_{\mathrm{pairwise}}$ a pairwise mono-channel dissimilarity measure that can be chosen among the wide range of pairwise measures conventionally used in image registration such as the sum of squared differences, or mutual information [105]. One multi-channel image is taken as fixed reference (here $\tilde{M}_1$). The other multi-channel image is known as the moving image (here $\tilde{M}_2$). The aim of pairwise multi-channel image registration is to minimise the following function:

$$\begin{aligned} \hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \; & \mathcal{D}_{\mathrm{PMC}}(\tilde{M}_{1,f}(\boldsymbol{x}), ..., \tilde{M}_{1,F}(\boldsymbol{x}), \\ & \tilde{M}_{2,1}(\boldsymbol{T}(\boldsymbol{x}, \boldsymbol{\mu})), ..., \tilde{M}_{2,F}(\boldsymbol{T}(\boldsymbol{x}, \boldsymbol{\mu}))) \end{aligned} \tag{5.3}$$

with $\boldsymbol{T}(\boldsymbol{x}, \boldsymbol{\mu})$ the registration transformation applied to the channels of the moving image $\tilde{M}_{2,1}, ..., \tilde{M}_{2,F}$, where $\boldsymbol{\mu}$ is a vector containing the parameters of the transformation, e.g. rotations and translations for a rigid transformation model, or B-spline coefficients for a free-form deformation model [118].

When $G > 2$ initially acquired images are considered, the pairwise multi-channel image registration scheme can still be used, but it should be applied $G-1$ times. One multi-channel image must be chosen as fixed reference image, e.g. $\tilde{M}_1$, to which the other images are registered. Disadvantages of such a scheme are that each independent registration is performed using only a fraction of the total available image information, and that the registration result depends on the choice of reference image [45].

5.2.3 Existing groupwise mono-channel registration scheme

Contrary to pairwise registration, groupwise registration allows to register multiple images in one optimisation procedure, without the need to select a fixed image space

reference [4, 55, 76, 94, 126, 150]. The aim of groupwise mono-channel registration is to simultaneously bring the $G \geq 2$ mono-channel images $M_g(\boldsymbol{x})$ of $\tilde{\mathcal{M}}$ to a mid-point space, by means of a transformation $\boldsymbol{T}(\boldsymbol{x}, \boldsymbol{\mu})$, where $\boldsymbol{\mu}$ is a vector containing the parameters $\boldsymbol{\mu}_g$ of the transformation $\boldsymbol{T}_g(\boldsymbol{x}, \boldsymbol{\mu}_g)$ that is applied to each image M_g. Let us denote $\mathcal{D}$ a measure quantifying the dissimilarity between all transformed images $M_g(\boldsymbol{T}_g(\boldsymbol{x}, \boldsymbol{\mu}_g))$. Groupwise mono-channel registration can then be formulated as the minimisation of the dissimilarity measure $\mathcal{D}$ with respect to $\boldsymbol{\mu}$:

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{D}(M_1(\boldsymbol{T}_1(\boldsymbol{x}, \boldsymbol{\mu}_1)), ..., M_G(\boldsymbol{T}_G(\boldsymbol{x}, \boldsymbol{\mu}_G))) \tag{5.4}$$

subject to:

$$\sum_{g=1}^{G} \boldsymbol{\mu}_g = \boldsymbol{0} \tag{5.5}$$

where the constraint of Equation (5.5) serves to define a mid-point space [10]. An attractive property of groupwise registrations is that they are fully symmetric (i.e. all images play the same role in the registration procedure), and are transitive consistent by definition.

5.2.4 Novel groupwise multi-channel image registration

The aim of the novel registration scheme that we propose is to solve multi-channel image registration by treating it as a groupwise registration problem. Simply extending the optimisation scheme presented in Equation 5.4 to channel images taken as individual images would not be satisfactory. Indeed, this would not take into account the knowledge that some images belong together, while others do not. The multi-channel nature of the images introduces an extra constraint: the channel images $\tilde{M}_{g,1}, ..., \tilde{M}_{g,F_g}$ of a given multi-channel image $\tilde{M}_g$ should be warped using the same transformation $\boldsymbol{T}_g(\boldsymbol{\mu}_g)$. We therefore incorporate a restricted transformation model within the standard groupwise registration approach of Equation (5.4). The new optimisation problem that we propose to solve is:

$$\begin{aligned} \hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{D}(&\tilde{M}_{1,1}(\boldsymbol{T}_1(\boldsymbol{\mu}_1)), ..., \tilde{M}_{1,F_1}(\boldsymbol{T}_1(\boldsymbol{\mu}_1)), ... \\ &\tilde{M}_{g,1}(\boldsymbol{T}_g(\boldsymbol{\mu}_g)), ..., \tilde{M}_{g,F_g}(\boldsymbol{T}_g(\boldsymbol{\mu}_g)), ... \\ &\tilde{M}_{G,1}(\boldsymbol{T}_G(\boldsymbol{\mu}_G)), ..., \tilde{M}_{G,F_G}(\boldsymbol{T}_G(\boldsymbol{\mu}_G))) \end{aligned} \tag{5.6}$$

This optimisation is subject to Equation (5.5). In this scheme, image dissimilarity $\mathcal{D}$ is assessed in an analogous way as in the existing groupwise mono-channel registration scheme (see Section 5.2.3), but the different optimisation procedure induces that the images are warped by groups.

This novel groupwise multi-channel image registration scheme allows to take into account correlations between any pair of channel images $\tilde{M}_{g,f}$, and not just the corresponding channels that have the same feature index f, which is a limitation of pairwise multi-channel registration (Equation (5.2)). In our framework, we consider all $\sum_{g=1}^{G} F_g$ images together. In this way, the information that is shared between

any pair of channel images can be taken into account, and not only the information shared between corresponding channels. Moreover, the scheme that we propose can be directly extended to datasets for which different numbers of channels are computed for each multi-channel image: in other words, our method does not suppose that there exists corresponding channel images, computed from each original image.

5.2.5 Choice of the dissimilarity measure

Multiple intensity-based groupwise dissimilarity metrics have been proposed and could be used within the novel multi-channel registration framework described in Section 5.2.4.

A particular case would be that the images of $\tilde{\mathcal{M}}$ are multi-channel and monomodal. In that case, choosing the groupwise dissimilarity measure based on the sum of variances, devised by Metz et al. [94] for images with similar intensities, would be appropriate.

In the most general case, however, the $\sum_{g=1}^{G} F_g$ images of $\tilde{\mathcal{M}}$ may have different contrasts. The groupwise dissimilarity measure to choose within the novel scheme should ideally handle such contrast differences. This is the case of the groupwise dissimilarity measures based on principal component analysis developed by Huizinga et al. [55], or based on mutual information as developed by Bhatia et al. [14], or of the total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ presented in Chapter 4.

The total correlation dissimilarity measure for groupwise registration, denoted $\mathcal{D}_{\mathrm{TC}}$) and fully described in Section 4.2, is one of the multivariate versions of mutual information. Experiments performed in Section 4.3 have shown that total correlation yields equal or better registration results than competing state-of-the-art methods on quantitative MRI datasets. Given the performances of total correlation on such non-monomodal datasets, we made the choice to select it as groupwise dissimilarity measure within the multi-channel and generic registration framework described in Section 5.2.4. This choice of dissimilarity measure was made to apply the novel scheme in a concrete case, but other choices in terms of dissimilarity measures and other registration components could have been made (see Section 5.5). The total correlation dissimilarity measure $\mathcal{D}_{\mathrm{TC}}$ is based on the computation of a correlation matrix $\mathbf{K}$, obtained from the images $M_1, ..., M_G$ in the existing groupwise mono-channel registration scheme, and obtained from the channel images $\tilde{M}_1, ..., \tilde{M}_G$ in the novel groupwise multi-channel registration scheme that we propose (see Section 5.2.4). Total correlation is computed as follows:

$$\mathcal{D}_{\mathrm{TC}} = \frac{1}{2} \sum_{j=1}^{G} \ln \lambda_j \tag{5.7}$$

with λ_j the j^{th} eigenvalue ($\lambda_j > \lambda_{j+1}$) of the correlation matrix $\mathbf{K}$ based on the image intensities of the images or channel images to register.

5.2.6 Optimisation methods and implementation details

The adaptive stochastic gradient descent (ASGD) of Klein et al. [64] was used as method for solving the optimisation problem described in Equation (5.6). This optimisation method randomly samples positions in the image space at each iteration to reduce computation time. For the ASGD optimisation, the derivative of the cost function has to be computed. To that purpose, we followed the work of van der Aa et al. [142] to differentiate the eigenvalue decomposition, similarly to what Huizinga et al. [55] did. Sampling is done off the voxel grid, which was shown to be necessary to reduce interpolation artefacts [65]. We chose to use an Euler parametrisation for rigid transformations (i.e. linear with 3 rotations and 3 translations, without scaling) and a B-spline model for non-rigid (i.e. non-linear) transformations.

A multi-resolution strategy was used: the images are Gaussian-blurred with a standard deviation that is decreased at each resolution level. With this procedure, the large deformations are corrected first, and the finer deformations are corrected in subsequent levels. Linear interpolation is used to interpolate the images during registration, which reduces computation time without significantly compromising registration accuracy [55], but cubic B-spline interpolation was used to produce the final registered images.

The registration framework that we propose supports the inclusion of regularisation terms in a trivial way [65]. However, we chose to leave these considerations out of the present work, to avoid the introduction of another hyperparameter (weight of the regularisation term). Instead, we keep the B-spline control point (when applicable) rather conservative, which intrinsically already provides some regularisation.

Many other choices for rigid (e.g. [149, 155]) and nonrigid (e.g. [8, 15, 29, 89, 114, 118, 145]) transformation models exist and could have been used in this study, but also when it comes to the optimisation methods (e.g. gradient descent [95], quasi-Newton [32], nonlinear conjugate gradient [28], Kiefer–Wolfowitz [62], simultaneous perturbation [125], Robbins–Monro [112], and evolution strategy [52]), different groupwise dissimilarity measures (see Section 5.2.3), multi-resolution strategies [79], and different regularisation terms (e.g. [39,115,127]). Comparing these different methods exhaustively is outside the scope of this chapter. What we propose is a generic framework that can be customised for different applications.

5.3 Experiments

The groupwise multi-channel registration scheme presented in Section 5.2.4 was evaluated on two multimodal imaging datasets, and compared to three other registration scenarios.

5.3.1 Registration scenarios

Four registration scenarios are compared. Each of them is detailed in this section and illustrated in Figure 5.1.

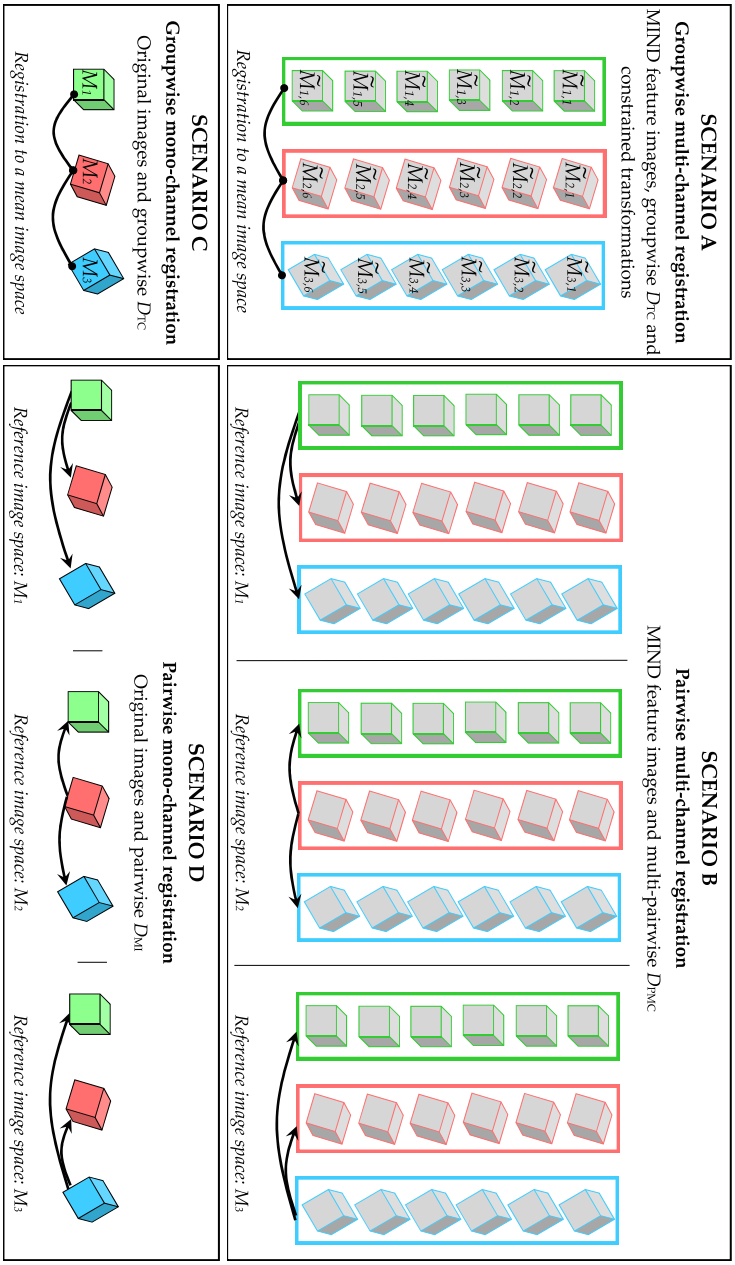


Figure 5.1: *Registration scenarios. Case with 3 original images M_1, M_2 and M_3.*

5.3.1.1 Scenario A – Groupwise multi-channel image registration

This scenario uses the novel groupwise multi-channel scheme introduced in Section 5.2.4, Equation (5.6). The feature images that we use are the modality independent neighbourhood descriptors (MIND) introduced by Heinrich et al. [53]. MIND feature images are obtained based on local variance and patch-based distances, using the following expression:

$$\tilde{M}_{g,f}(\boldsymbol{x}) = \frac{1}{n} \exp\left(-\frac{D_p(M_g, \boldsymbol{x}, \boldsymbol{x} + \boldsymbol{r}_f)}{V(M_g, \boldsymbol{x})}\right) \tag{5.8}$$

where D_p is a patch-based distance, V an estimation of local variance, and n a normalisation constant. F spatial search vectors $\boldsymbol{r}_1, ..., \boldsymbol{r}_F$ serve to compute F MIND feature images $\tilde{M}_{g,1}, ..., \tilde{M}_{g,F}$ for each original image M_g. For a given search vector $\boldsymbol{r}_f$, the MIND images obtained from different original images $\tilde{M}_{1,f}, ..., \tilde{M}_{G,f}$ have a quite similar appearance [50, 53], as shown in Figure 5.2. However, this is not the case when different $\boldsymbol{r}_f$ are considered. As justified in Section 5.2.5, total correlation $\mathcal{D}_{\mathrm{TC}}$ is taken as groupwise dissimilarity measure in that scenario.

5.3.1.2 Scenario B – Pairwise multi-channel image registration

In this second scenario, we use the existing pairwise multi-channel scheme described in Equation (5.3). MIND feature images are used, like in scenario A, to build the multi-channel set of images that have to be registered. MIND feature images obtained for a given index f have similar intensity distributions, irrespective of the modality of the original image. The pairwise multi-channel dissimilarity measure that we choose to consider in this scenario is mutual information $\mathcal{D}_{\mathrm{MI}}$ [105, 149]:

$$\mathcal{D}_{\mathrm{PMC}}(\tilde{\mathcal{M}}) = \sum_{f=1}^{F} \mathcal{D}_{\mathrm{MI}}\left(\tilde{M}_{1,f}, \tilde{M}_{2,f}\right) \tag{5.9}$$

with M_1, M_2 two images with N samples.

5.3.1.3 Scenario C – Groupwise mono-channel registration

Scenario C consists of applying existing groupwise mono-channel registration to the original images, without using any feature image. The registration scheme is the one corresponding to Equation (5.4), using $\mathcal{D}_{\mathrm{TC}}$ as dissimilarity measure.

5.3.1.4 Scenario D – Pairwise mono-channel registration

In scenario D, we apply pairwise mono-channel registration based on mutual information $\mathcal{D}_{\mathrm{MI}}$ to the original images.

5.3.1.5 Additional groupwise scenarios

To provide insight into the influence of the choice of the groupwise dissimilarity measure on the registration results, experiments similar to scenario A and C were repeated, but with another dissimilarity measure than $\mathcal{D}_{\mathrm{TC}}$. Huizinga et al. [55] previously designed a groupwise dissimilarity measure based on principal component analysis, denoted $\mathcal{D}_{\mathrm{PCA2}}$, the expression of which is close to $\mathcal{D}_{\mathrm{TC}}$. In their study, Huizinga et al. [55] concluded that the registration results obtained with $\mathcal{D}_{\mathrm{PCA2}}$ were similar to or better than results obtained with other state-of-the-art techniques, which is why we also performed results with that dissimilarity measure. We additionally performed the same experiments with Wachinger and Navab's [150] dissimilarity measure based on accumulated pairwise estimates, $\mathcal{D}_{\mathrm{APE}}$, which is not based on mutual information and not closely related to $\mathcal{D}_{\mathrm{TC}}$.

5.3.2 Experiment 1 – Head and neck multimodal dataset

Twenty-two patients with a tumour in the head and neck region [40,41,146,147] were scanned for radiotherapy and hyperthermia treatment planning [37, 104]. Approval was obtained from the institutional review board in regard to our study (number: METC-2010-318). A multimodal dataset was acquired for each patient, consisting of a T1-weighted magnetic resonance image (MRI), a T2-weighted MRI, and a computed tomography (CT) image. The CT images were acquired using a Siemens scanner (Somatom Sensation Open, Siemens), with a voxel size of $0.98\times0.98\times2.50$ mm^3 for 21 patients, and $1.27\times1.27\times2.50$ mm^3 for the remaining patient. The T1 and T2-weighted MR images were acquired on a 1.5 T scanner (Optima MR450w, General Electric Healthcare). Half of the patients had MR images with a voxel size of $0.49\times0.49\times3.00$ mm^3, the other half of $0.68\times0.68\times3.00$ mm^3. The out-of-plane field of view (FOV) was 20 cm centered at the location of the tumour, yielding slightly different FOV values among patients. T1 and T2 images were corrected for intensity inhomogeneity, using the N3 method [123] for the T1 images, and the built-in algorithm (surface coil intensity correction) of the MR scanner for the T2 images. Radiotherapy immobilisation masks were used [41] both for the acquisition of the MR and CT images. For 12 patients (subset α), the immobilisation masks covered the head, the neck and the shoulders. For this first subset, the MR images were acquired using 6-channel flex coils allowing for the same immobilisation position in both the MR and CT acquisitions. For the remaining ten patients (subset β), smaller immobilisation masks were used, covering only the anterior parts of the head and neck region. For this second subset, the use of head coils made it impossible to impose the exact same patient position in the CT and MR acquisitions.

5.3.2.1 Image preparation

In the current implementation of the method, the groupwise multi-channel and mono-channel schemes operate provided that the images to register have equal voxel sizes. To ensure that this is the case, the T2 and CT images were first coarsely registered to the T1 image using a rigid transformation and pairwise mutual information.

The images were subsequently misaligned using artificially generated Euler transformations. The image data was resampled only once for these operations. These transformations were randomly generated using uniform distributions (translations $\pm$10 mm, rotations $\pm$0.25 rad). To generate the MIND feature images, we chose a Gaussian weighting of σ = 2, and a six-neighbourhood spatial region for the search vectors $\boldsymbol{r}_1, ..., \boldsymbol{r}_F$, resulting in F = 6 feature images for each acquired image [50, 53]. Examples of such images are shown in Figure 5.2a.

5.3.2.2 Registration settings

For each patient, the original images after image preparation are denoted M_1 (the T1-weighted MRI), M_2 (the T2-weighted MRI) and M_3 (the CT image).

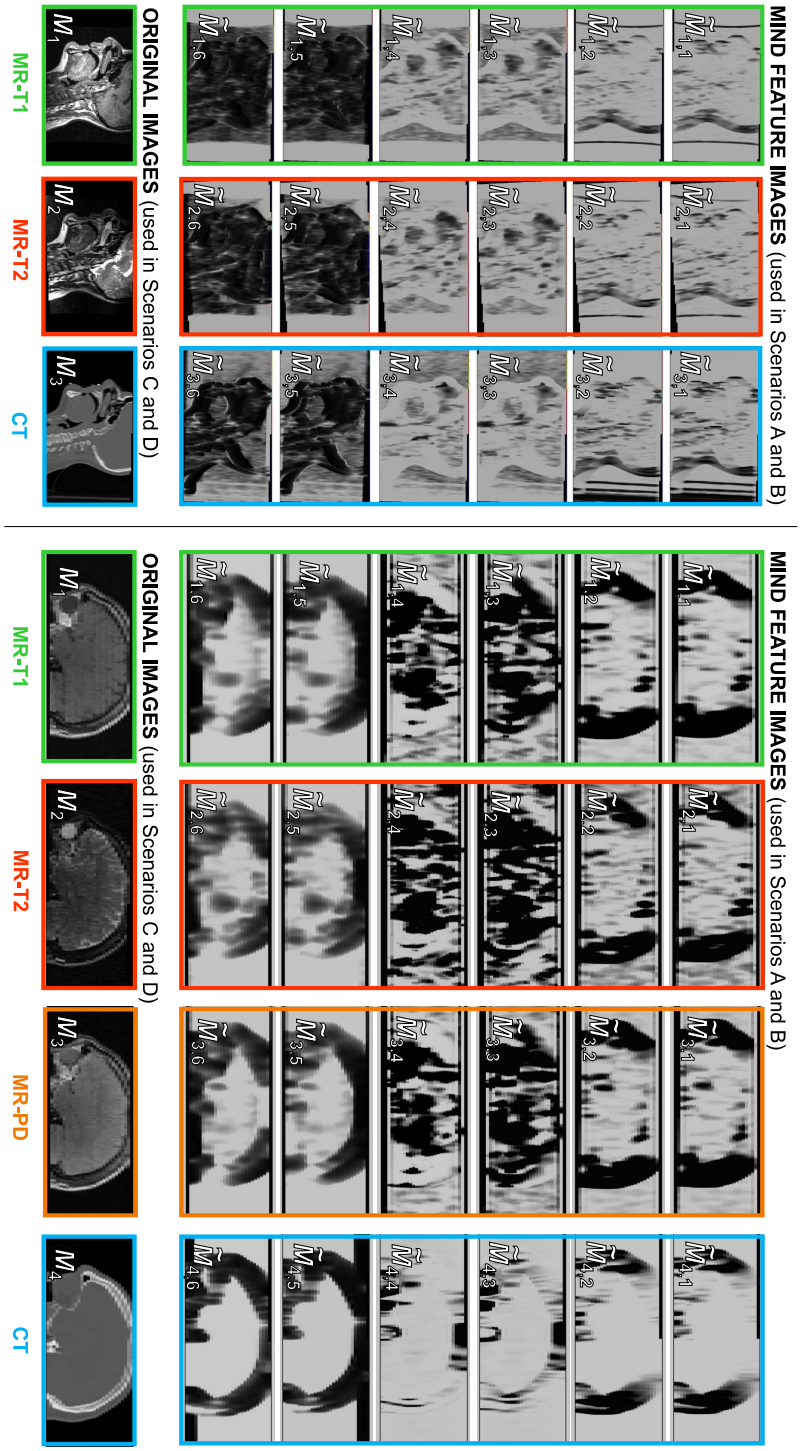
The four image registration scenarios described in Section 5.3.1. are then applied to the image datasets. Registrations were applied in two consecutive steps. Firstly, with a rigid transformation model based on Euler transformations, and secondly with a deformable model based on B-spline transformations [118] with control point spacing of 100 mm, to account for possible non-rigid misalignments due to different patient bulk positioning or organ positioning (especially in the neck region) between the scans. The Euler transformations were taken as starting points for the registrations based on B-splines. Initial trial-and-error experiments suggested that the value of 100 mm is sufficient to compensate for the deformations in the head and neck images. Three resolutions of 1,000 iterations each with a smoothing σ of 8, 4 and 2 voxels were used in both the Euler and B-spline cases. For the pairwise mono-channel registrations based on mutual information, the number of bins that we selected was 32. A B-spline Parzen windowing approach was used to estimate the probability density function for mutual information [133]. Furthermore, registration masks for the head and neck region were delineated for each patient for all modalities in order to prevent the influence on registration of artefacts such as ghosting for the MR images, and the presence of the immobilisation. Considering that the MIND feature images show noisy backgrounds, the registration masks seemed to be particularly necessary for the registration of these MIND images (Figure 5.2).

5.3.2.3 Registration evaluation

For each patient, an expert placed n = 19$\pm$2 corresponding landmarks for all imaging modalities. Registration accuracy was evaluated by computing the following target registration error (TRE) between the landmarks of the three modalities:

$$\begin{aligned} \mathrm{TRE} &= \frac{1}{3}\left(d_{\mathrm{T1-T2}} + d_{\mathrm{T1-CT}} + d_{\mathrm{T2-CT}}\right) \\ &\text{with } d_{\boldsymbol{p}-\boldsymbol{q}} = \frac{1}{n}\sum_{\mathrm{i}=1}^{n} ||\boldsymbol{p}_i - \boldsymbol{q}_i|| \end{aligned} \tag{5.10}$$

with $\boldsymbol{p}_i$ and $\boldsymbol{q}_i$ the landmark coordinates, $||\,.\,||$ the Euclidean distance, and $d_{\boldsymbol{p}-\boldsymbol{q}}$ the average landmark distance between the modalities $\boldsymbol{p}$ and $\boldsymbol{q}$. The head and neck dataset allows to study the influence of the choice of reference image: T1, T2 or CT. In



(a) Head and neck dataset

(b) RIRE dataset

Figure 5.2: *Examples of MIND feature images and original images for the head and neck dataset (a) and for the RIRE dataset (b).*

the pairwise multi-channel case (scenario B) and in the pairwise mono-channel case (scenario D), the landmarks were directly propagated to the three image reference spaces of T1, T2 and CT. In the groupwise cases (scenarios A and C), no reference space is chosen during registration (all images are brought to an average space), but the evaluation of registration accuracy can be done in the image spaces of the original images. To that purpose, the landmarks were brought to the average image space, and were subsequently propagated to each of the original image spaces T1, T2 or CT by using the inverses of the transformations obtained by the groupwise registrations [94]. Wilcoxon signed-rank tests were used to assess significance of comparisons of TRE values with respect to scenario A. Significance was considered for $p < 0.01$. Intra-rater variability (IRV) was assessed by repeating the placement of landmarks one month after the first placement session. This operation was done for the T2 and CT images before the artificial misalignments were applied. Landmarks were not annotated on T1 because of the natural alignment of the T1 and T2 images at the moment of acquisition. For each patient, the IRV was computed using $\text{IRV} = \frac{1}{2}\left(d_{\text{T2}-\text{T2}^\diamond} + d_{\text{CT}-\text{CT}^\diamond}\right)$, $\diamond$ referring to the second landmark delineation session.

5.3.3 Experiment 2 – RIRE multimodal dataset

This experiment focuses on multimodal images of the publicly available Retrospective image registration evaluation (RIRE) project [155]. Out of the 18 patient datasets available on the website of the RIRE project[1], we selected the 12 datasets including at least the three following modalities: CT, MR-T1 and MR-T2. Seven of these 12 datasets also included a proton density-weighted MR image (MR-PD). PET images were not considered in this study because of the presence of halo artefacts. CT images had a voxel size between $0.40\times0.40\times3.00\,\text{mm}^3$ and $0.65\times0.65\times4.00\,\text{mm}^3$, and MR images had a voxel size between $0.82\times0.82\times3.00\,\text{mm}^3$ and $1.25\times1.25\times4.00\,\text{mm}^3$. One of the patients was not included because the field of view of the MRIs was much smaller than for the CT image.

5.3.3.1 Image preparation

The initial step was to resample the T1, T2 and PD images (when present) to the image space of the CT image. MIND images are computed using the same settings as described in Section 5.3.2.1. Examples of such images are shown in Figure 5.2b.

5.3.3.2 Registration settings

The registration scenarios that we considered are the same as for Experiment 1. Only the CT image was taken as fixed reference space in scenarios B and D due to the specific requirements of the online evaluation tool. Registrations were applied in two consecutive steps. Firstly, with a translation transformation model, and secondly with an Euler transformation model. The translation transformations were taken as starting points for the registrations based on Euler. Non-rigid transformations were not considered for RIRE datasets, following the guidelines of the RIRE online

[1] http://www.insight-journal.org/rire/

platform. In a similar fashion to the head and neck dataset, registration masks were used to exclude the background during registration. The registration parameters that were used here are identical to those of Experiment 1 (Section 5.3.2.2).

5.3.3.3 Registration evaluation

The registrations were evaluated by uploading lists of points to the website of the RIRE project. The ground truth is made available in an indirect manner only: users cannot get direct access to the ground truth, and instead have to upload points transformed using the registration results. In contrast to the head and neck dataset, the ground truth is here only known with respect to the CT images, which is why the influence of the choice of reference image could not be studied in the framework of this experiment. Wilcoxon signed-rank test were used to assess significance of comparisons of TRE values with respect to scenario A.

5.3.4 Experiment 3 – Groupwise multi-channel registration for multi-channel images with different numbers of channels

As mentioned in Section 5.2, the groupwise multi-channel scheme described in this chapter offers the possibility to register datasets of multi-channel images with different numbers of channels. For the head and neck dataset, we performed groupwise multi-channel registration based on total correlation $\mathcal{D}_{\mathrm{TC}}$ with all MIND channels for T1 ($\tilde{M}_{1,1}, ..., \tilde{M}_{1,6}$), with 4 MIND channels for T2 ($\tilde{M}_{2,1}, ..., \tilde{M}_{2,4}$), and with 5 MIND channels for CT ($\tilde{M}_{3,1}, ..., \tilde{M}_{3,5}$). The registration settings are the same as those described in Experiment 1 (Section 5.3.2.2).

5.3.5 Implementation

All registration methods used in this study were implemented as part of the open source package elastix [65]. For the experiments, image manipulations were performed using Python (version 2.7.3) with packages NumPy 1.6.2, SciPy 0.11.0, pydicom 0.9.7, and NiBabel 1.3.0. To obtain the MIND feature images, we directly used the MATLAB code made available by its authors[2]. All computations were done on a Linux machine (64 GB of memory, 8 cores, 2412 MHz).

5.4 Results

5.4.1 Results on Experiment 1 – Head and neck dataset

Target registration errors (TRE) obtained for the head and neck dataset are presented in Table 5.1 for the four registration scenarios and for the three reference spaces T1, T2 and CT. Results obtained for the groupwise scenarios with $\mathcal{D}_{\mathrm{PCA2}}$ and $\mathcal{D}_{\mathrm{APE}}$ are presented in Section 5.7. *Average 1* and *STD 1* are obtained based on the TRE values

[2]http://www.ibme.ox.ac.uk/research/biomedia/julia-schnabel/Software

that correspond to a given scenario and a given reference space. *Average 2* and *STD 2* are obtained based on all TRE values for a given scenario.

Focusing first on the *Average 2* values, it is observed that scenario C resulted in gross registration failures, which was the case with $\mathcal{D}_{\mathrm{TC}}$, with $\mathcal{D}_{\mathrm{PCA2}}$ and with $\mathcal{D}_{\mathrm{APE}}$. The lowest registration errors were obtained with scenario B, which was significantly lower than for scenarios A and D. Similar patterns are observed when focusing on *Average 1* values. A discordant note is the fact that the *Average 1* TRE values for scenario D with CT as reference image were substantially larger than the TRE values with the T1 and T2 as reference images. In addition, the difference in *Average 1* TRE values between scenario A and D with CT as reference image was statistically significant.

Table 5.1: *Experiment 1 - TRE [mm] for each reference space T1, T2 and CT. Bold values of Average 1 and Average 2 signal distributions that are significantly different from scenario A (underlined values). For Average 1, the comparisons were performed per reference space. For Average 2, all data was considered together.*

			Ⓐ Groupwise multi-channel			Ⓑ Pairwise multi-channel			Ⓒ Groupwise mono-channel			Ⓓ Pairwise mono-channel			
	Patient	Mis.	T1	T2	CT	T1	T2	CT	T1	T2	CT	T1	T2	CT	IRV
α : Large radiotherapy mask	1	24.7	2.2	2.2	2.2	2.2	2.1	2.1	106.8	113.1	108.2	1.7	2.0	2.4	0.8
	2	14.9	2.3	2.3	2.2	2.1	2.2	2.1	101.9	110.8	105.0	2.1	2.2	2.6	1.1
	3	22.2	2.0	2.0	2.0	1.7	2.0	1.7	8.4	8.3	9.4	1.7	1.7	2.0	0.8
	4	22.9	2.3	2.3	2.3	1.9	2.2	2.2	115.8	112.8	126.7	1.9	2.2	2.6	0.8
	5	15.8	2.6	2.6	2.6	2.4	2.3	2.5	8.2	8.3	8.2	2.5	2.5	3.6	0.8
	6	17.2	2.5	2.5	2.5	2.3	2.4	2.4	97.2	99.4	97.6	2.3	2.3	3.2	0.9
	7	21.9	2.0	2.0	2.0	2.1	1.8	1.9	11.0	10.8	12.2	1.8	2.0	2.4	0.8
	8	18.5	2.1	2.1	2.1	2.0	2.0	2.1	25.2	25.1	25.1	1.9	2.1	3.0	0.8
	9	19.9	1.9	1.9	1.9	1.7	1.8	1.9	94.0	108.1	108.6	1.6	1.7	2.5	1.1
	10	19.3	2.4	2.4	2.3	2.2	2.1	2.1	107.0	109.7	102.3	2.2	2.0	2.8	0.8
	11	15.3	2.5	2.5	2.5	3.5	3.0	2.4	123.0	135.5	124.6	3.4	2.4	2.3	1.1
	12	19.8	2.3	2.2	2.2	2.2	2.5	2.2	65.5	62.7	61.8	2.4	1.9	3.4	1.0
β : Small radiotherapy mask	13	26.4	5.2	5.2	5.0	4.2	3.7	4.5	10.1	10.0	10.6	4.0	4.9	4.7	1.1
	14	21.7	2.7	2.7	2.7	2.5	2.7	2.9	12.1	12.1	12.8	2.6	2.6	2.8	0.8
	15	19.9	2.7	2.7	2.7	2.5	2.5	2.6	6.5	6.4	6.2	2.7	2.9	3.1	0.7
	16	23.3	3.4	3.4	3.4	2.4	3.0	2.9	9.0	9.1	8.7	2.8	3.0	4.0	0.7
	17	16.3	3.8	3.8	3.9	3.3	3.9	3.5	7.3	7.3	7.1	4.3	4.7	5.2	1.1
	18	15.3	4.0	4.0	4.1	4.0	3.7	4.1	10.2	10.0	11.0	4.2	4.4	5.2	0.9
	19	10.9	3.3	3.3	3.4	2.9	3.2	3.6	14.8	14.8	14.8	3.9	4.4	6.1	1.1
	20	22.5	2.5	2.5	2.5	1.8	2.2	2.2	41.9	42.0	38.9	1.8	1.8	12.7	1.1
	21	20.7	3.1	3.1	3.2	3.2	3.1	3.3	19.5	19.5	19.0	2.6	2.9	2.9	0.9
	22	21.1	3.2	3.2	3.2	2.7	2.7	2.6	4.8	4.8	5.0	2.9	3.4	3.1	1.2
Average 1		**19.6**	2.8	2.8	2.8	**2.5**	**2.6**	**2.6**	**45.5**	**47.3**	**46.5**	2.6	2.7	**3.8**	0.9
STD 1		3.6	0.8	0.8	0.8	0.7	0.6	0.7	44.0	46.8	45.7	0.8	1.0	2.2	0.2
Variability		-	7×10^{-3}			3×10^{-2}			0.8			0.5			-
Average 2		**19.6**	2.8			**2.6**			**46.4**			3.0			0.9
STD 2		3.6	0.8			0.7			45.5			1.6			0.2

Comparing the TRE values between the reference spaces T1, T2 and CT allows to assess the influence of the choice of reference image on registration accuracy. Stan-

dard deviation values were therefore computed between the three average TRE values obtained in each image reference space to measure the registration consistency across the three reference spaces (see Section 5.3.2.3). These standard deviations are denoted *Variability* in Table 5.1, and are computed based on the three corresponding values of *Average 1*. The lowest variability was obtained for scenario A. When inspecting the individual TRE values for each subject, this becomes even more apparent.

The results of Table 5.1 clearly indicate that the TRE values obtained for subset α (large immobilisation mask) are lower than for subset β (small immobilisation mask). This can be explained by the fact that the immobilisation mask and gradient coils used for subset β did not allow a similar positioning of the patient between the CT and MR acquisitions. In particular, rotations of the neck region may make the registration more challenging. The intra-rater variability averaged over all patients is 0.9 mm, which is about three times lower than the best *Average 2* TRE.

Table 5.2: *Experiment 2 - TRE [mm]. Bold values signal distributions that are significantly different from scenario A (underlined value).*

Patient	Misaligned case	Ⓐ Groupwise multi-channel	Ⓑ Pairwise multi-channel	Ⓒ Groupwise mono-channel	Ⓓ Pairwise mono-channel
001	25.1	2.6	2.1	12.9	4.1
002	36.1	1.1	3.3	33.3	3.1
005	38.9	2.3	8.9	16.6	2.5
006	41.0	2.5	12.9	10.6	6.8
007	25.1	1.6	5.9	7.3	2.9
101	14.6	4.7	1.9	58.0	2.5
102	11.6	2.9	2.6	15.2	2.0
105	28.1	7.5	4.5	78.5	8.6
106	23.3	5.8	3.9	6.2	2.3
107	20.9	2.3	2.7	4.7	1.5
108	27.6	3.9	2.8	69.6	1.9
109	27.2	2.3	2.7	85.8	1.7
Average	**26.6**	3.3	**4.3**	**33.2**	3.3
STD	8.5	1.8	3.1	29.6	2.1

5.4.2 Results on Experiment 2 – RIRE dataset

For the RIRE dataset, the registration accuracy results are presented in Table 5.2. Results obtained for the additional scenarios with $\mathcal{D}_{\mathrm{PCA2}}$ and $\mathcal{D}_{\mathrm{APE}}$ are presented in Section 5.7. The average TRE before registration is 26.6 mm. After registration, the average TRE is 3.3 mm for scenario A, 4.3 mm for scenario B, 33.2 mm for scenario C and 3.3 mm for scenario D. Contrary to Experiment 1, the best registration results

were here obtained with groupwise multi-channel registration (scenario A), and with pairwise mono-channel registration (scenario D).

5.4.3 Results on Experiment 3 – Groupwise multi-channel registration for multi-channel images with different numbers of channels

In this experiment, we considered the images obtained from patients 1 to 12 of the head and neck dataset. Using the set of images $\{\tilde{M}_{1,1}, ..., \tilde{M}_{1,6}, \tilde{M}_{2,1}, ..., \tilde{M}_{2,4}, \tilde{M}_{3,1}, ..., \tilde{M}_{3,5}\}$, which includes different number of channels for each multi-channel image, we obtained an average TRE of 2.3 mm, which is only slightly worse than the TRE of 2.2 mm obtained in the original experiment with all channels.

5.5 Discussion

In this chapter, we presented a novel groupwise multi-channel image registration scheme. This scheme can be combined with a wide range of groupwise dissimilarity measures and sets of feature images. As a proof of concept, we applied this groupwise multi-channel registration scheme to MIND feature images, and used total correlation $\mathcal{D}_{\mathrm{TC}}$ as groupwise dissimilarity measure.

Groupwise multi-channel image registration (scenario A) was compared to a pairwise multi-channel registration scheme (scenario B). The first main theoretical advantage of the groupwise multi-channel scheme is that it is fully symmetric, and therefore has transitive consistency by definition. Additionally, groupwise multi-channel image registration is directly extensible to datasets for which different numbers of feature images are computed for each image, which is not the case for the pairwise multi-channel dissimilarity measure. Results on two multimodal datasets (head and neck and RIRE) indicate that groupwise multi-channel registration achieves similar registration accuracy as pairwise multi-channel registration, in terms of TRE (scenario A yielded better TRE results than scenario B for the RIRE dataset, but slightly worse for the head and neck dataset). In terms of variability, the experiments on the head and neck dataset confirm the interest of the symmetric groupwise formulation, in the sense that TRE results depend more on the reference space in the pairwise methods than in the groupwise multi-channel approach. The poor variability figures obtained in the groupwise mono-channel case is attributed to the complete registration failure obtained in that case.

These two multi-channel schemes (scenarios A and B) were also compared to mono-channel scenarios based on the original images only, either in a groupwise manner (scenario C), or in a pairwise manner (scenario D). The fact that scenario C results in a registration failure, while scenario A did not, illustrates the potential of the novel scheme that we propose. The results indicate that even when groupwise mono-channel registration with a given dissimilarity measure results in registration failure, there is a possibility to improve the registration results by computing appropriate feature images from the original images (e.g. using MIND feature images) and

applying the proposed groupwise multi-channel registration with the same dissimilarity measure. Besides, the experiment on the head and neck dataset indicates that groupwise multi-channel registration (scenario A) is much less sensitive towards the choice of reference space than pairwise mono-channel registration (scenario D), which verifies one of the advantages of groupwise registration with respect to pairwise registration.

Groupwise total correlation was previously applied to monomodal and quantitative MR images (in Chapter 4) and did not require the use of feature images to obtain results similar or better than pairwise mutual information [48, 55]. The fact that the groupwise total correlation dissimilarity measure requires feature images like MIND images for the registration of multimodal datasets might be explained by the fact that such feature images are more similar to one another, and therefore easier to register (which indeed was the prime motivation for developing MIND [53]). The approximation of total correlation used for the sake of computational feasibility are apparently too drastic in the case of true multimodal images, while it worked on multi-parametric images as shown in [49]. MIND pre-processing allows to make the images sufficiently similar so that they can be registered using total correlation.

The focus of our study was to present a novel multi-channel groupwise registration technique, and does not consist of a comparison between multiple pre-processing algorithms. However, many other pre-processing methods, like ALOST [81] could have been tested. Many choices would have been possible with respect to the transformation model, regularisation, optimisation, pre-processing features and dissimilarity measures (see Section 5.2.6), in particular. Comparing various such components is outside of the scope of this chapter. Our key contribution, i.e. the formulation of multi-channel image registration as a groupwise registration problem, is independent of these choices. We do not claim that the proposed groupwise multi-channel method is per se better than mono-channel registration. Some dissimilarity measures might be better in coping with certain differences in intensities than others, and it depends on the image characteristics which dissimilarity measure is optimal. For pairwise registration, mutual information is a proven robust method that works in many multimodal registration scenarios. Yet, it was shown that using additional features yields improved performances in certain applications [53, 78, 128]. For groupwise situations, no analogous approach to mutual information has been presented yet, so there the use of multiple feature channels, as proposed in our study, is a worthwhile option to consider. However, we cannot claim that this multi-channel approach will lead *in general* to better performance.

5.6 Conclusion

In this chapter, we described a scheme for the groupwise registration of multi-channel images. We showed that the registration of multiple sets of feature images can be solved effectively with a groupwise multi-channel registration method, using previously proposed intensity dissimilarity measures suitable for multiparametric imaging data, in combination with a restricted transformation model that assigns a single transformation to all feature images that belong together. In this way, the

shared information between all feature images of all images is taken into account, the number of feature images in each set is flexible, the registration is unbiased (i.e. there is no need to choose a reference frame), and the approach naturally scales to scenarios with more than two sets of feature images.

5.7 Supplementary materials

The tables presented in this section provide quantitative TRE results in the groupwise multi-channel registration scheme applied to the MIND images, and on the groupwise mono-channel scheme applied to the acquired head and neck and RIRE images. These results, based on two groupwise dissimilarity measures ($\mathcal{D}_{\mathrm{PCA2}}$ developed by Huizinga et al. [55], and $\mathcal{D}_{\mathrm{APE}}$ developed by Wachinger and Navab [150]), complement those presented in Tables 5.1 and 5.2.

Table 5.3: *Experiment 1 – TRE [mm] for each reference space T1, T2 and CT. Bold values of Average 1 and Average 2 signal distributions that are significantly different from scenario A (underlined values in Table 5.1). For Average 1, the comparisons were performed per reference space. For Average 2, all data was considered together.*

			Groupwise multi-channel with $\mathcal{D}_{\mathrm{PCA2}}$			Groupwise mono-channel with $\mathcal{D}_{\mathrm{PCA2}}$			Groupwise multi-channel with $\mathcal{D}_{\mathrm{APE}}$			Groupwise mono-channel with $\mathcal{D}_{\mathrm{APE}}$			
	Patient	Mis.	T1	T2	CT	T1	T2	CT	T1	T2	CT	T1	T2	CT	IRV
α : Large radiotherapy mask	1	24.7	2.2	2.2	2.2	104.3	107.6	104.5	2.5	2.5	2.4	106.7	114.1	108.2	0.8
	2	14.9	2.2	2.2	2.1	102.1	108.9	103.9	2.4	2.4	2.3	102.9	110.1	104.8	1.1
	3	22.2	2.0	1.9	1.9	6.2	6.2	5.9	2.1	2.1	2.0	8.3	8.2	9.5	0.8
	4	22.9	2.2	2.2	2.2	100.5	103.1	94.9	2.5	2.6	2.5	149.6	149.6	135.7	0.8
	5	15.8	2.6	2.5	2.6	8.3	8.3	8.3	2.7	2.7	2.6	11.0	11.0	11.2	0.8
	6	17.2	2.6	2.6	2.6	92.5	96.0	93.8	2.5	2.6	2.5	106.3	106.5	106.0	0.9
	7	21.9	2.0	2.0	2.0	10.8	10.6	12.2	2.2	2.2	2.1	11.0	10.8	12.2	0.8
	8	18.5	2.1	2.1	2.1	34.0	33.8	32.9	2.1	2.1	2.1	35.5	35.2	34.3	0.8
	9	19.9	1.9	1.9	1.9	100.0	107.1	96.9	2.2	2.2	2.1	100.8	113.8	99.7	1.1
	10	19.3	2.5	2.5	2.4	101.6	111.9	104.3	2.3	2.3	2.2	102.5	105.9	99.3	0.8
	11	15.3	2.4	2.4	2.4	120.6	137.4	131.6	2.9	2.9	2.8	118.4	137.0	129.0	1.1
	12	19.8	2.3	2.3	2.2	74.4	71.3	72.5	2.3	2.4	2.3	95.0	89.7	99.1	1.0
β : Small radiotherapy mask	13	26.4	5.4	5.4	5.2	9.8	9.3	10.3	4.4	4.4	4.4	9.8	9.7	10.2	1.1
	14	21.7	2.7	2.7	2.7	44.6	44.6	41.0	2.7	2.7	2.7	40.9	40.9	38.0	0.8
	15	19.9	2.9	2.9	2.9	6.2	6.1	5.9	2.4	2.5	2.4	6.9	6.9	6.6	0.7
	16	23.3	3.2	3.2	3.2	9.5	9.5	9.4	2.6	2.6	2.7	16.9	17.0	17.9	0.7
	17	16.3	3.9	3.9	3.9	7.4	7.4	7.3	3.6	3.6	3.6	7.6	7.7	7.5	1.1
	18	15.3	4.1	4.0	4.2	10.4	10.2	11.2	3.7	3.7	3.8	8.1	7.9	8.7	0.9
	19	10.9	3.3	3.3	3.4	12.7	12.8	12.7	3.1	3.1	3.2	12.9	13.0	13.0	1.1
	20	22.5	2.6	2.6	2.5	51.2	51.8	47.7	2.1	2.1	2.1	54.5	54.3	49.2	1.1
	21	20.7	3.2	3.2	3.3	44.2	43.8	41.0	3.1	3.1	3.2	43.8	43.4	39.8	0.9
	22	21.1	3.3	3.2	3.3	116.4	116.0	122.1	3.0	3.0	3.0	4.7	4.7	4.9	1.2
Average 1		**19.6**	2.8	2.8	2.8	**53.1**	**55.2**	**53.2**	2.7	2.7	2.7	**52.5**	**54.4**	**52.0**	0.9
STD 1		3.6	0.8	0.8	0.8	43.9	46.9	45.0	0.6	0.6	0.6	47.7	50.5	47.2	0.2
Variability		-	1×10^{-2}			1.2			1×10^{-2}			1.2			-
Average 2		**19.6**	2.8			**53.8**			2.7			**53.0**			0.9
STD 2		3.6	0.8			44.6			0.6			47.7			0.2

Table 5.4: *Experiment 2 - TRE [mm]. Bold values signal distributions that are significantly different from scenario A (underlined values in Table 5.2).*

Patient	Misaligned case	Groupwise multi-channel with $\mathcal{D}_{\mathrm{PCA2}}$	Groupwise mono-channel with $\mathcal{D}_{\mathrm{PCA2}}$	Groupwise multi-channel with $\mathcal{D}_{\mathrm{APE}}$	Groupwise mono-channel with $\mathcal{D}_{\mathrm{APE}}$
001	25.1	4.0	13.1	2.6	12.1
002	36.1	2.1	32.1	1.3	49.6
005	38.9	4.0	18.2	112.1	16.1
006	41.0	5.3	10.7	99.9	10.7
007	25.1	1.3	7.3	1.3	7.8
101	14.6	5.1	60.1	4.4	52.0
102	11.6	3.6	15.9	2.4	15.1
105	28.1	6.9	152.8	6.0	10.5
106	23.3	5.5	6.4	5.6	6.7
107	20.9	2.2	4.6	2.4	4.9
108	27.6	3.6	9.6	3.1	12.1
109	27.2	2.1	81.0	2.4	82.0
Average	**26.6**	**3.8**	**34.3**	**20.3**	**23.3**
STD	8.5	1.8	4.0	29.6	2.1

General discussion and conclusion

Image registration has become an essential tool in medical imaging. It is used in a wide scope of applications to support diagnosis, therapy planning, or guidance of interventions and surgery. The number of registration methods that have been developed, especially in the past two decades, has followed the increasing diversity of image datasets.

Devising a registration method for a specific application may be pursued following two main approaches. The first consists of arranging existing image registration methods, adapting them if necessary, selecting the adequate registration components (e.g. dissimilarity measure, transformation model, optimisation method), and establishing a satisfactory parameter tuning. We followed this approach in **Chapters 2** and **3**, in which we developed advanced registration pipelines adapted to specific applications. This approach is however not necessarily suitable when specific registration problems have to be solved, or when novel types of medical image datasets are considered. The second approach consists of devising new registration components (e.g. new dissimilarity measures), or even entirely new registration schemes. Sometimes, such components have already been presented, but have not been brought to the world of medical image registration. In this thesis, we designed novel registration components: a total correlation groupwise dissimilarity measure (**Chapter 4**), and a registration scheme for the registration of datasets of multi-channel images (**Chapter 5**).

6.1 Main contributions of this thesis

In this thesis, we devised registration techniques for quantitative imaging and multi-channel images. Both the methodological and clinical contributions related to these techniques are highlighted in this section.

6.1.1 Methodological contributions

In **Chapter 2**, we developed an image registration pipeline for the registration of diffusion-weighted MR (DW-MR) images acquired from the abdominal area of healthy volunteers. We showed that image registration improves the reproducibility of the apparent diffusion coefficient (ADC), a quantitative parameter obtained from multiple DW-MR images. Our registration pipeline compensates for misalignments both within each acquired image, and between multiple images. The curve fitting step used to extract the ADC parameters, based on maximum likelihood, makes it possible to compute a measure of uncertainty of the fit (the Fréchet-Cramér-Rao lower bound) at each voxel location, such that not only the ADC value but also the accuracy with which it is extracted is also derived. We additionally showed that trivial operations such as data blurring, interpolation, or averaging of multiple non-registered images yields ADC values that are biased and less reproducible than when image registration is applied. On the datasets we considered, the ADC values directly extracted from the acquired images are overestimated by 30% with respect to the case in which our image registration pipeline is applied. This is in line with the results obtained in the context of a collaborative follow-up effort [109] in which the work of Chapter 2 was used as reference approach.

Chapter 3 focuses on cystic fibrosis (CF) patients with respiratory tract exacerbation (RTE). DW-MR images were acquired using a MR image acquisition sequence related to the one used in Chapter 2, resulting in datasets with lacking spatial correspondence within and between the images. In Chapter 3, we adapted the registration pipeline designed in Chapter 2 to the concrete clinical application of cystic fibrosis, knowing the verified advantages of that pipeline in terms of the values and reproducibility of the quantitative parameters extracted from the DW-MR images. In Chapter 3, we considered a two-step curve fitting method based on maximum likelihood for extracting three quantitative imaging parameters from the DW-MR images using the intravoxel incoherent motion (IVIM) model. Another methodological contribution of this chapter is a weighted local smoothing method based on the Fréchet-Cramér-Rao lower bounds computed thanks to the maximum likelihood fitting step. This weighted local smoothing technique allows to reduce the impact that unreliably fitted values have on the parameters extracted from the quantitative images.

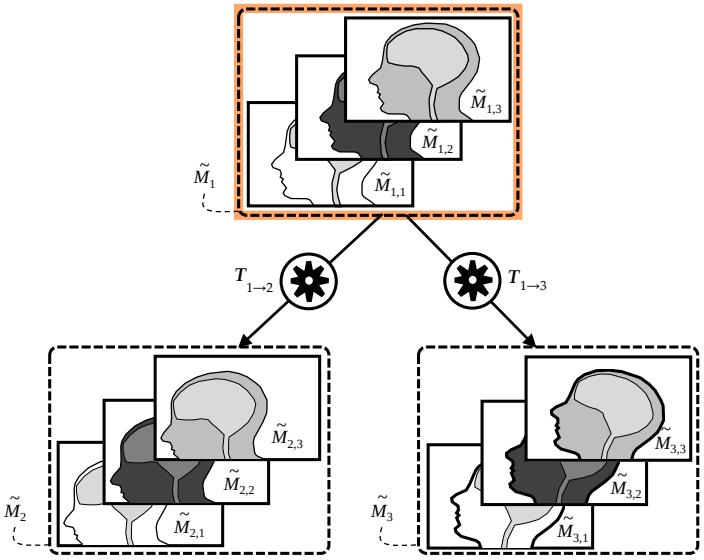
Chapter 4 makes it possible to apply one of the multivariate versions of mutual information in the context of groupwise image registration. Our interest in multivariate dissimilarity measures based on mutual information lies in the fact that pairwise mutual information is successfully used in common practice in a large number of applications. Moreover, our interest is justified by the lack of groupwise equivalents or extensions used for image registration. The groupwise dissimilarity measure considered in Chapter 4, called total correlation, was selected based on both theoretical and implementation advantages with respect to other multivariate versions of mutual information. Besides, we approximated total correlation for implementation purposes. The results show that groupwise total correlation based on total correlation yields better registration accuracy results than pairwise registration based on mutual information, on quantitative MR and dynamic imaging datasets.

In **Chapter 5**, we developed a new groupwise scheme for the registration of datasets of multi-channel images, i.e. images for which multiple sub-images (called channels) are available. Contrary to the existing pairwise multi-channel registration techniques (Figure 6.1a), the groupwise multi-channel technique that we proposed in Chapter 5 (Figure 6.1b) allows to register all feature images simultaneously, does not require the selection of a reference image space, and assigns a single transformation to all feature images that belong together. In addition, our groupwise multi-channel scheme does not require the existence of the same number of feature channels per feature image. We showed that even if a groupwise dissimilarity measure does not make it possible to register a set of images, our scheme offers the technical possibility to increase registration accuracy by considering feature images based on the original images to register. Our scheme is modular in terms of registration components such as the groupwise dissimilarity measure, the number of feature images, the transformation model, and can be tailored per application.

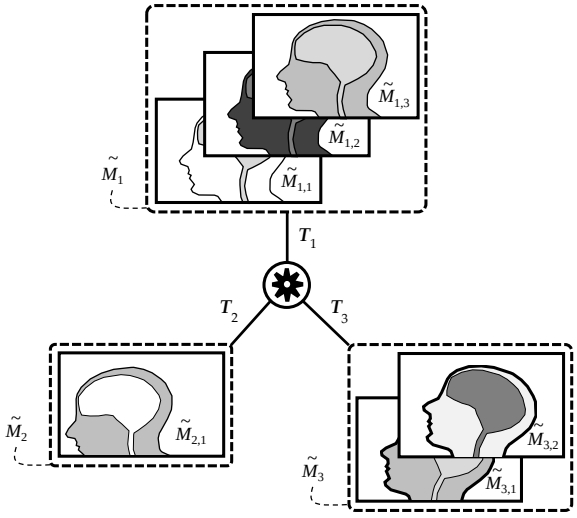
6.1.2 Clinically-oriented contributions

In **Chapter 2**, we showed that ensuring spatial coherence between images is an essential step to derive reliable quantitative measurements from them. One of the aims in quantitative imaging consists of devising quantitative imaging biomarkers that are technically validated (i.e. they can be measured reliably) and qualified (i.e. they can be linked to clinical implications), as exposed by Waterton et al. [154]. Chapter 2 verified that image registration plays an essential role in ensuring the technical validity of quantitative ADC measurements obtained from multiple DW-MR images, by increasing the reproducibility with which the ADCs are obtained and by providing accurate quantification. ADC is a quantitative measurement that has been considered in the context of pathophysiological processes, such as cancer, that are known to have an impact on diffusion properties in tissues.

In **Chapter 3**, we extracted quantitative measurements based on the intravoxel incoherent motion (IVIM) theory, which contrary to the ADC technique, allows to quantify molecular diffusion and perfusion in a separate manner. Chapter 3 considers DW-MR images acquired from CF patients. Respiratory tract exacerbation (RTE) scores is a clinical measure developed to distinguish the presence from the absence of pulmonary exacerbations [116]. This standardised measure allows comparisons between different studies, and is established based on a questionnaire filled by physicians, including questions about observed symptoms, physical examination, patient history and pulmonary function test results. RTE scores are global, functional and qualitative measures that are computed for a given patient at a given time point. In Chapter 3, we proposed to monitor RTE using the intravoxel incoherent motion (IVIM) theory, introduced by Le Bihan et al. [73]. In particular, we show that the IVIM quantitative parameter D^*, characterising blood perfusion, is a promising quantitative parameter for capturing treatment response in CF patients affected by RTE. This quantitative parameter is extracted from multiple DW-MR images using a curve fitting technique. The image registration pipeline developed and tested on analogous volunteer datasets in Chapter 2 was used in Chapter 3 to ensure spa-



(a) Pairwise registration of multi-channel images.
The orange box indicates that M_1 is taken as reference image



(b) Groupwise registration of multi-channel images
proposed in Chapter 5

Figure 6.1: *Example of pairwise (a) and groupwise (b) methods for the registration of multi-channel datasets of medical images. The illustrated groupwise technique was proposed in Chapter 5 of this thesis.*

tial correspondence between all DW-MR images, and therefore ensure a reliable extraction of the quantitative parameter D^*. Contrary to the funcational qualitative RTE scores, D^* can be assessed at any spatial coordinate, and can therefore offer a per-lesion characterisation of pulmonary exacerbation. Another added value of D^* is that it is obtained based on diffusion-weighted MR images, which allows to perform multiple scanning sessions for patient monitoring, while this would not be possible with techniques based on computed tomography (CT) due to radiation dose limitations [69].

In terms of clinical applications, the total correlation dissimilarity measure developed in **Chapter 4** can be applied to datasets of quantitative MR images or to datasets of dynamic images. In this chapter, we considered applications in the myocardium, carotid artery wall, spleen, brain parenchyma, pancreas, and in the lung, based on various imaging modalities: CT, T_1-weighted MRI, DW-MRI, and DCE-MRI. The groupwise dissimilarity measure based on total correlation that we developed is a tool that can be used to register and, if applicable, extract more reliable quantitative measurements from datasets of medical images that can be described by a low-dimensional signal model.

The groupwise multi-channel registration method developed in **Chapter 5** can be used for the registration of multi-channel medical images. Such images may be created by applying various sorts of filtering to images, like we did through the use of the MIND feature images. Our groupwise multi-channel method also opens new promising perspectives for studies based on datasets of multi-channel images generated by the imaging modality themselves, such as in dual-echo MR imaging [59, 93], in contrast-enhanced ultrasound imaging [2], or in dual-energy CT [91].

6.2 Perspectives for future research

A further research step for the quantitative imaging experiments of Chapters 2 and 3 would be to register all even and odd subvolumes simultaneously in a single groupwise registration procedure, avoiding the sequential registration pipeline that we considered in these chapters. One of the challenges in groupwise image registration is the choice of dissimilarity measure, especially when multimodal images have to be registered. Even though we showed in Chapter 4 that groupwise total correlation yields registration results that are comparable to or better than pairwise mutual information in the case of quantitative MRI and dynamic imaging, further research needs to be performed for the groupwise registration of any kind of multimodal image dataset, as shown in Chapter 5. One possible axis for future research works could be to refine total correlation, or elaborate other groupwise dissimilarity measures that would be as widely applicable as mutual information is in pairwise image registration. For the time being, we obtained promising multi-modal registration results using total correlation by applying it to sets of feature images obtained from the original images to register. To that end, we used the multi-channel groupwise registration technique of Chapter 5.

6.3 Conclusion

This thesis describes and evaluates advanced medical image registration methods for quantitative imaging and multi-channel images. We proposed registration methods that improve the reliability and reproducibility of quantitative measurements extracted in a voxelwise manner from multiple medical images, and applied them to multiple medical imaging modalities focusing on different anatomical areas of interest or pathologies. We also designed a novel registration scheme that allows the registration of images for which multiple channels exist. The image registration techniques that we developed are modular in the sense that the registration components can be tailored for a given application. Our work sheds a particular light on groupwise registration techniques, as these techniques might get an increasing interest in medical imaging in the coming years. In addition to bringing about advances in registration accuracy, groupwise registration could be an elegant and efficient way of decomplexifying image registration pipelines in a time marked by an inflation in the numbers and types of medical images acquired in the clinic.

Bibliography

[1] N. A. Ahmed & D. V. Gokhale, "Entropy expressions and their estimators for multivariate distributions", *IEEE Trans Inf Theory*, vol. 35, no. 3, pp. 688–692, 1989.

[2] Z. Akkus, D. D. Carvalho, S. C. van den Oord, A. F. Schinkel, W. J. Niessen, N. de Jong, A. F. van der Steen, S. Klein, J. G. Bosch, "Fully automated carotid plaque segmentation in combined contrast-enhanced and B-mode ultrasound", *Ultrasound Med Biol*, vol. 41, no. 2, pp. 517–531, 2015.

[3] R. Amin, M. Charron, L. Grinblat, A. Shammas, H. Grasemann, K. Graniel, P. Ciet, H. Tiddens, F. Ratjen, "Cystic fibrosis: detecting changes in airway inflammation with FDG PET/CT", *Radiology*, vol. 264, no. 3, pp. 868–75, 2012.

[4] J. L. R. Andersson & S. Skare, "A model-based method for retrospective correction of geometric distortions in diffusion-weighted EPI.", *Neuroimage*, vol. 16, no. 1, pp. 177–199, 2002.

[5] A. Andreou, D.-M. Koh, D. J. Collins, M. Blackledge, T. Wallace, M. O. Leach, M. R. Orton, "Measurement reproducibility of perfusion fraction and pseudodiffusion coefficient derived by intravoxel incoherent motion diffusion-weighted MR imaging in normal liver and metastases", *Eur Radiol*, vol. 23, no. 2, pp. 428–434, 2013.

[6] L. R. Arlinghaus, E. B. Welch, A. B. Chakravarthy, L. Xu, J. S. Farley, V. G. Abramson, A. M. Grau, M. C. Kelley, I. A. Mayer, J. A. Means-Powell, I. M. Meszoely, J. C. Gore, T. E. Yankeelov, "Motion correction in diffusion-weighted MRI of the breast at 3T", *J Magn Reson Imaging*, vol. 33, no. 6, pp. 693–704, 2011.

[7] R. Attariwala & W. Picker, "Whole body MRI: improved lesion detection and characterization with diffusion weighted techniques", *J Magn Reson Im*, vol. 38, no. 2, pp. 253–268, 2013.

[8] B. Avants & J. C. Gee, "Geodesic estimation for large deformation anatomical shape averaging and interpolation", *NeuroImage*, vol. 23, pp. S139–S150, 2004.

[9] B. B. Avants, C. L. Epstein, M. Grossman, J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain", *Med Imag Anal*, vol. 12, no. 1, pp. 26–41, 2008.

[10] S. Balci, P. Golland, M. Shenton, M. Wells, "Free-form B-spline deformation model for groupwise registration", in *MICCAI*, pp. 23–30 (Statistical Registration Workshop), 2007.

[11] A. J. Bell, "Co-information lattice", in *4th International Symposium on Independent Component Analysis and Blind Source Separation*, pp. 921–926, 2003.

[12] R. Bellman, *Adaptive control processes: a guided tour*, Princeton University Press, New Jersey, 1961.

[13] M. A. Bernstein, K. F. King, X. J. Zhou, *Handbook of MRI pulse sequences*, Elsevier Inc, Burlington, 2004.

[14] K. K. Bhatia, J. Hajnal, A. Hammers, D. Rueckert, "Similarity metrics for groupwise non-rigid registration", in *MICCAI*, vol. 2, pp. 544–552, 2007.

[15] F. L. Bookstein, "Principal warps: thin-plate splines and the decomposition of deformations", 1989.

[16] E. E. Bron, J. van Tiel, H. Smit, D. H. J. Poot, W. J. Niessen, G. P. Krestin, H. Weinans, E. H. G. Oei, G. Kotek, S. Klein, "Image registration improves human knee cartilage T1 mapping with delayed gadolinium-enhanced MRI of cartilage (dGEMRIC)", *Eur Radiol*, vol. 23, no. 1, pp. 246–252, 2013.

[17] M. Bruegel, K. Holzapfel, J. Gaa, K. Woertler, S. Waldt, B. Kiefer, A. Stemmer, C. Ganter, E. J. Rummeny, "Characterization of focal liver lesions by ADC measurements using a respiratory triggered diffusion-weighted single-shot echo-planar MR imaging technique", *Eur Radiol*, vol. 18, no. 3, pp. 477–485, 2008.

[18] R. Castillo, E. Castillo, R. Guerra, V. E. Johnson, T. McPhail, a. K. Garg, T. Guerrero, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets", *Phys Med Biol*, vol. 54, no. 7, pp. 1849–1870, 2009.

[19] G. B. Chavhan, P. S. Babyn, S. S. Vasanawala, "Abdominal MR imaging in children: motion compensation, sequence optimization, and protocol organization", *Radiographics*, vol. 33, no. 3, pp. 703–719, 2013.

[20] M. Chen, A. Carass, A. Jog, J. Lee, S. Roy, J. L. Prince, "Cross contrast multi-channel image registration using image synthesis for MR brain images", *Med Image Anal*, vol. 36, pp. 2–14, 2017.

[21] P. Ciet, S. Bertolo, M. Ros, E. R. Andrinopoulou, V. Tavano, F. Lucca, T. Freiweier, G. P. Krestin, H. A. W. M. Tiddens, G. Morana, "Detection and monitoring of lung inflammation in cystic fibrosis during respiratory tract exacerbation using diffusion-weighted magnetic resonance imaging", *Eur Respir J*, vol. 50, no. 1, pp. 1–9, 2016.

[22] P. Ciet, G. Serra, E. R. Andrinopoulou, S. Bertolo, M. Ros, C. Catalano, S. Colagrande, H. A. W. M. Tiddens, G. Morana, "Diffusion weighted imaging in cystic fibrosis disease: beyond morphological imaging", *Eur Radiol*, vol. 26, no. 11, pp. 3830–3839, 2016.

[23] P. Ciet, G. Serra, S. Bertolo, S. Spronk, M. Ros, F. Fraioli, S. Quattrucci, M. B. Assael, C. Catalano, F. Pomerri, H. A. W. M. Tiddens, G. Morana, "Assessment of CF lung disease using motion corrected PROPELLER MRI: a comparison with CT", *Eur Radiol*, vol. 26, no. 3, pp. 780–787, 2016.

[24] B. F. Coolen, D. H. J. Poot, M. I. Liem, L. P. Smits, S. Gao, G. Kotek, S. Klein, A. J. Nederveen, "Three-dimensional quantitative T1 and T2 mapping of the carotid artery: sequence design and in vivo feasibility", *Magn Reson Med*, vol. 75, no. 3, pp. 1–10, 2015.

[25] T. M. Cover & J. A. Thomas, *Elements of information theory*, Hoboken, New Jersey, 2005.

[26] M. D. Craene & S. K. Warfield, "Multi-subject registration for unbiased statistical atlas construction", in *MICCAI*, pp. 655–662, 2004.

[27] W. R. Crum, T. Hartkens, D. L. G. Hill, "Non-rigid image registration: theory and practice", *Brit J Radiol*, vol. 77, pp. S140–S153, 2004.

[28] Y.-H. Dai, "A family of hybrid conjugate gradient methods for unconstrained optimization", *Math Comput*, vol. 72, no. 243, pp. 1317–1329, 2003.

[29] M. H. Davis, A. Khotanzad, D. P. Flamig, S. E. Harms, "A physics-based coordinate transformation for 3-D image matching", *IEEE T Med Imaging*, vol. 16, no. 3, pp. 317–328, 1997.

[30] N. De Geeter, G. Crevecoeur, L. Dupré, W. Van Hecke, A. Leemans, "A DTI-based model for TMS using the independent impedance method with frequency-dependent tissue parameters", *Phys Med Biol*, vol. 57, no. 8, pp. 2169–2188, 2012.

[31] Y. Deng, X. Li, Y. Lei, C. Liang, Z. Liu, "Use of diffusion-weighted magnetic resonance imaging to distinguish between lung cancer and focal inflammatory lesions: a comparison of intravoxel incoherent motion derived parameters and apparent diffusion coefficient", *Acta Radiol*, vol. 57, no. 11, pp. 1310–1317, 2015.

[32] J. E. Dennis, Jr., J. J. Moré, J. E. Dennis, J. J. Moré, "Quasi-Newton methods, motivation and theory", *SIAM Review*, vol. 19, no. 1, pp. 46–89, 1977.

[33] H. Dijkstra, P. Baron, P. Kappert, M. Oudkerk, P. E. Sijens, "Effects of microperfusion in hepatic diffusion weighted imaging", *Eur Radiol*, vol. 22, no. 4, pp. 891–899, 2012.

[34] G. Dournes, F. Menut, J. Macey, M. Fayon, J. F. Chateil, M. Salel, O. Corneloup, M. Montaudon, P. Berger, F. Laurent, "Lung morphology assessment of cystic fibrosis using MRI with ultra-short echo time at submillimeter spatial resolution", *Eur Radiol*, vol. 26, no. 11, pp. 3811–3820, 2016.

[35] M. Eichinger, C.-P. Heussel, H.-U. Kauczor, H. Tiddens, M. Puderbach, "Computed tomography and magnetic resonance imaging in cystic fibrosis lung disease", *J Magn Reson Im*, vol. 32, no. 6, pp. 1370–1378, 2010.

[36] U. Euler & G. Liljestrand, "Observations on the pulmonary arterial blood pressure in the cat", *Acta Physiol Scand*, vol. 12, no. 1937, pp. 301–320, 1946.

[37] P. Evans, "Anatomical imaging for radiotherapy", *Phys Med Biol*, vol. 53, no. 12, pp. 151–191, 2008.

[38] R. Failo, P. A. Wielopolski, H. A. W. M. Tiddens, W. C. J. Hop, R. P. Mucelli, M. H. Lequin, "Lung morphology assessment using MRI: a robust ultra-short TR/TE 2D steady state free precession sequence used in cystic fibrosis patients", *Magn Reson Med*, vol. 61, no. 2, pp. 299–306, 2009.

[39] B. Fischer & J. Modersitzki, "A unified approach to fast image registration and a new curvature based registration technique", *Linear Algebra Appl*, vol. 380, pp. 107–124, 2004.

[40] V. Fortunati, R. F. Verhaart, F. Angeloni, A. van der Lugt, W. J. Niessen, J. F. Veenland, M. M. Paulides, T. van Walsum, "Feasibility of multimodal deformable registration for head and neck tumor treatment planning", *Int J Radiat Oncol*, vol. 90, no. 1, pp. 85–93, 2014.

[41] V. Fortunati, R. F. Verhaart, G. M. Verduijn, A. Van Der Lugt, F. Angeloni, W. J. Niessen, J. F. Veenland, M. M. Paulides, T. Van Walsum, "MRI integration into treatment planning of head and neck tumors: can patient immobilization be avoided?", *Radiother Oncol*, vol. 115, no. 2, pp. 191–194, 2015.

[42] N. Fujima, D. Yoshida, T. Sakashita, A. Homma, A. Tsukahara, K. K. Tha, K. Kudo, H. Shirato, "Intravoxel incoherent motion diffusion-weighted imaging in head and neck squamous cell carcinoma: assessment of perfusion-related parameters compared to dynamic contrast-enhanced MRI", *Magn Reson Imaging*, vol. 32, no. 10, pp. 1206–1213, 2014.

[43] D. J. Galas, N. A. Sakhanenko, A. Skupin, T. Ignac, "Describing the complexity of systems: multivariable set complexity and the information basis of systems biology", *J Comput Biol*, vol. 21, no. 2, pp. 118–140, 2014.

[44] P. Gaviani, R. B. Schwartz, E. T. Hedley-Whyte, K. L. Ligon, A. Robicsek, P. Schaefer, J. W. Henson, "Diffusion-weighted imaging of fungal cerebral infection", *Am J Neuroradiol*, vol. 26, no. 5, pp. 1115–1121, 2005.

[45] X. Geng, G. E. Christensen, H. Gu, T. J. Ross, Y. Yang, "Implicit reference-based group-wise image registration and its application to structural and functional MRI", *NeuroImage*, vol. 47, no. 4, pp. 1341–1351, 2009.

[46] C. H. Goss & J. L. Burns, "Exacerbations in cystic fibrosis - 1: epidemiology and pathogenesis", *Thorax*, vol. 62, no. 4, pp. 360–367, 2007.

[47] V. Gulani, F. Calamante, F. G. Shellock, E. Kanal, S. B. Reeder, "Gadolinium deposition in the brain: summary of evidence and recommendations", *Lancet Neurol*, vol. 16, no. 7, pp. 564–570, 2017.

[48] J.-M. Guyader, L. Bernardin, N. H. M. Douglas, D. H. J. Poot, W. J. Niessen, S. Klein, "Influence of image registration on apparent diffusion coefficient images computed from free-breathing diffusion MR images of the abdomen.", *J Magn Reson Im*, vol. 42, no. 2, pp. 315–330, 2015.

[49] J.-M. Guyader, W. Huizinga, V. Fortunati, D. H. Poot, M. V. Kranenburg, J. F. Veenland, M. M. Paulides, W. J. Niessen, S. Klein, "Total correlation-based groupwise image registration for quantitative MRI", in *WBIR*, pp. 186–193, 2016.

[50] J.-M. Guyader, W. Huizinga, V. Fortunati, J. F. Veenland, M. M. Paulides, W. J. Niessen, S. Klein, "Groupwise image registration of multimodal head-and-neck images", in *ISBI*, pp. 730–733, 2015.

[51] T. S. Han, "Nonnegative entropy measures of multivariate symmetric correlations", *Inform Control*, vol. 36, no. 2, pp. 133–156, 1978.

[52] N. Hansen & A. Ostermeier, "Completely derandomized self-adaptation in evolution strategies", *Evol Comput*, vol. 9, no. 2, pp. 159–195, 2001.

[53] M. P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. V. Gleeson, S. M. Brady, J. A. Schnabel, "MIND: modality independent neighbourhood descriptor for multi-modal deformable registration", *Med Image Anal*, vol. 16, no. 7, pp. 1423–1435, 2012.

[54] M. P. Heinrich, B. W. Papiez, J. A. Schnabel, H. Handels, "Multispectral image registration based on local canonical correlation analysis", *MICCAI*, vol. 17, pp. 202–209, 2014.

[55] W. Huizinga, D. H. J. Poot, J. Guyader, R. Klaassen, B. F. Coolen, M. V. Kranenburg, "PCA-based groupwise image registration for quantitative MRI", *Med Image Anal*, vol. 29, pp. 65–78, 2016.

[56] L. Ibáñez, W. Schroeder, L. Ng, J. Cates, *The ITK software guide*, second edi ed., Kitware Inc, New York, 2005.

[57] M. K. Ivancevic, T. C. Kwee, T. Takahara, T. Ogino, H. K. Hussain, P. S. Liu, T. L. Chenevert, "Diffusion-weighted MR imaging of the liver at 3.0 Tesla using TRacking Only Navigator echo (TRON): a feasibility study", *J Magn Reson Imaging*, vol. 30, no. 5, pp. 1027–1033, 2009.

[58] N. P. Jerome, M. R. Orton, J. a. D'Arcy, D. J. Collins, D.-M. Koh, M. O. Leach, "Comparison of free-breathing with navigator-controlled acquisition regimes in abdominal diffusion-weighted magnetic resonance images: effect on ADC and IVIM statistics", *J Magn Reson Imaging*, vol. 39, no. 1, pp. 235–240, 2014.

[59] B. A. Johnson, E. K. Fram, B. P. Drayer, B. L. Dean, P. J. Keller, R. Jacobowitz, "Evaluation of shared-view acquisition using repeated echoes (SHARE): a dual-echo fast spin-echo MR technique", *Am J Neuroradiol*, vol. 15, no. 4, pp. 667–673, 1994.

[60] S. Joshi, B. Davis, M. Jomier, G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy", *NeuroImage*, vol. 23, no. Suppl. 1, pp. S151–S160, 2004.

[61] H. Kandpal, R. Sharma, K. S. Madhusudhan, K. S. Kapoor, "Respiratory-triggered versus breath-hold diffusion-weighted MRI of liver lesions: comparison of image quality and apparent diffusion coefficient values", *Am J Roentgenol*, vol. 192, no. 4, pp. 915–922, 2009.

[62] J. Kiefer & J. Wolfowitz, "Stochastic estimation of the maximum of a regression function", *Ann Math Stat*, vol. 23, no. 3, pp. 462–466, 1952.

[63] R. Klaassen, O. Gurney-Champion, E. ter Voert, A. Heerschap, M. Bijlsma, M. Besselink, G. van Tienhoven, J. Nio, P. C., J. Wilmink, H. van Laarhoven, A. Nederveen, "Motion correction of high temporal 3T dynamic contrast enhanced MRI of pancreatic cancer - preliminary results", in *ISMRM*, p. 3667 (Poster), 2014.

[64] S. Klein, J. Pluim, M. Staring, M. Viergever, "Adaptive stochastic gradient descent optimisation for image registration", *Int J Comput Vision*, vol. 81, no. 3, pp. 227–239, 2009.

[65] S. Klein, M. Staring, K. Murphy, M. Viergever, J. Pluim, "Elastix: a toolbox for intensity-based medical image registration", *IEEE T Med Imaging*, vol. 29, no. 1, pp. 196–205, 2010.

[66] D.-M. Koh & D. J. Collins, "Diffusion-weighted MRI in the body: applications and challenges in oncology", *Am J Roentgenol*, vol. 188, no. 6, pp. 1622–1635, 2007.

[67] D.-M. Koh, D. J. Collins, M. R. Orton, "Intravoxel incoherent motion in body diffusion-weighted MRI: reality and challenges", *Am J Roentgenol*, vol. 196, no. 6, pp. 1351–1361, 2011.

[68] D.-M. Koh, T. Takahara, Y. Imai, D. J. Collins, "Practical aspects of assessing tumors using clinical diffusion-weighted imaging in the body", *Magn Res Med Sci*, vol. 6, no. 4, pp. 211–224, 2007.

[69] W. Kuo, P. Ciet, H. A. W. M. Tiddens, W. Zhang, R. P. Guillerman, M. Van Straten, "Monitoring cystic fibrosis lung disease by computed tomography: radiation risk in perspective", *Am J Respir Crit Care*, vol. 189, no. 11, pp. 1328–1336, 2014.

[70] T. C. Kwee, T. Takahara, D.-m. Koh, R. A. J. Nievelstein, P. R. Luijten, "Comparison and reproducibility of ADC measurements in breathhold, respiratory MR imaging of the liver", *J Magn Reson Imaging*, vol. 28, no. 5, pp. 1141–1148, 2008.

[71] N. E. Larsen, S. Haack, L. P. S. Larsen, E. M. Pedersen, "Quantitative liver ADC measurements using diffusion-weighted MRI at 3 Tesla: evaluation of reproducibility and perfusion dependence using different techniques for respiratory compensation", *Magn Reson Mater Phy*, vol. 26, no. 5, pp. 431–442, 2013.

[72] D. Le Bihan, "Intravoxel incoherent motion perfusion MR imaging: a wake-up call", *Radiology*, vol. 249, no. 3, pp. 748–752, 2008.

[73] D. Le Bihan, E. Breton, D. Lallemand, M.-L. Aubin, J. Vignaud, M. Laval-Jeantet, "Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging", *Radiology*, vol. 168, no. 2, pp. 497–505, 1988.

[74] D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanie, M. Laval-Jeantet, "MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders", *Radiology*, vol. 161, no. 2, pp. 401–407, 1986.

[75] D. Le Bihan, R. Turner, J. Macfall, "Effects of intravoxel incoherent motions (IVIM) in steady-state free precession (SSFP) imaging: application to molecular diffusion imaging", *Magn Reson Med*, vol. 337, pp. 324–337, 1989.

[76] E. Learned-Miller, "Data Driven Image Manifolds through Continous Joint Alignment", *IEEE T Pattern Anal*, vol. 28, no. 2, pp. 236–250, 2006.

[77] A. Leemans, J. Sijbers, S. De Backer, E. Vandervliet, P. Parizel, "Multiscale white matter fiber tract coregistration: a new feature-based approach to align diffusion tensor data", *Magn Reson Med*, vol. 55, no. 6, pp. 1414–1423, 2006.

[78] P. A. Legg, P. L. Rosin, D. Marshall, J. E. Morgan, "A robust solution to multi-modal image registration by combining mutual information with multi-scale derivatives", in *MICCAI*, pp. 616–623, 2009.

[79] H. Lester & S. R. Arridge, "A survey of hierarchical non-linear medical image registration", *Pattern Recogn*, vol. 32, no. 1, pp. 129–149, 1999.

[80] Z. Li, D. Mahapatra, J. Tielbeek, J. Stoker, L. van Vliet, F. Vos, "Image registration based on autocorrelation of local structure", *IEEE T Med Imaging*, vol. 35, no. 1, pp. 63–75, 2016.

[81] Z. Li, D. Mahapatra, J. A. W. Tielbeek, J. Stoker, L. J. van Vliet, F. M. Vos, "Image registration based on autocorrelation of local structure", *IEEE T Med Imaging*, vol. 35, no. 1, pp. 63–75, 2016.

[82] J. Liau, J. Lee, M. E. Schroeder, C. B. Sirlin, M. Bydder, "Cardiac motion in diffusion-weighted MRI of the liver: artifact and a method of correction", *J Magn Reson Imaging*, vol. 35, no. 2, pp. 318–327, 2012.

[83] H. Liu, Y. Liu, T. Yu, N. Ye, "Usefulness of diffusion-weighted MR imaging in the evaluation of pulmonary lesions", *Eur Radiol*, vol. 20, no. 4, pp. 807–815, 2010.

[84] M. Loeve, W. C. J. Hop, M. De Bruijne, P. T. W. Van Hal, P. Robinson, M. L. Aitken, J. D. Dodd, H. A. W. M. Tiddens, "Chest computed tomography scores are predictive of survival in patients with cystic fibrosis awaiting lung transplantation", *Am J Respir Crit Care*, vol. 185, no. 10, pp. 1096–1103, 2012.

[85] A. Luciani, A. Vignaud, M. Cavet, "Liver cirrhosis: intravoxel incoherent motion MR imaging - Pilot study", *Radiology*, vol. 249, no. 3, pp. 891–899, 2008.

[86] A. Luna, J. Sanchez-Gonzalez, P. Caro, "Diffusion-weighted imaging of the chest", *Magn Reson Imaging Clin N Am*, vol. 19, no. 1, pp. 69–94, 2011.

[87] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, P. Suetens, "Image registration by maximization of mutual information", *IEEE T Med Imaging*, vol. 16, no. 2, pp. 187–198, 1997.

[88] J. B. A. Maintz & M. A. Viergever, "A survey of medical image registration", *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[89] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, W. Eubank, "PET-CT image registration in the chest using free-form deformations", *IEEE T Med Imaging*, vol. 22, no. 1, pp. 120–128, 2003.

[90] Y. Mazaheri, R. K. G. Do, A. Shukla-Dave, "Motion correction of multi-b-value diffusion-weighted imaging in the liver", *Acad Radiol*, vol. 19, no. 12, pp. 1573–1580, 2012.

[91] C. H. McCollough, S. Leng, L. Yu, J. G. Fletcher, "Dual- and multi-energy CT: principles, technical approaches, and clinical applications", *Radiology*, vol. 276, no. 3, pp. 637–653, 2015.

[92] W. J. McGill, "Multivariate information transmission", *Psychometrika*, vol. 19, no. 2, pp. 317–325, 1954.

[93] E. M. Merkle & R. C. Nelson, "Dual gradient-echo phase hepatic MR imaging: a useful tool for evaluating more than fatty infiltration or fatty sparing", *Radiographics*, vol. 26, no. 5, pp. 1409–1418, 2006.

[94] C. Metz, S. Klein, M. Schaap, T. van Walsum, W. Niessen, "Nonrigid registration of dynamic medical imaging data using nD + t B-splines and a groupwise optimization approach", *Med Image Anal*, vol. 15, no. 2, pp. 238–249, 2011.

[95] M. Milanič, W. Jia, J. S. Nelson, B. Majaron, *Numerical optimization of sequential cryogen spray cooling and laser irradiation for improved therapy of port wine stain*, vol. 43, 2011.

[96] J. Modersitzki, *Numerical methods for image registration*, Oxford Science Publications, Oxford, 2004.

[97] T. Moritani, J. Kim, A. A. Capizzano, P. Kirby, J. Kademian, Y. Sato, "Pyogenic and non-pyogenic spinal infections: Emphasis on diffusion-weighted imaging for the detection of abscesses and pus collections", *Brit J Radiol*, vol. 87, no. 1041, 2014.

[98] N. Murata, K. Murata, L. F. Gonzalez-Cuyar, K. R. Maravilla, "Gadolinium tissue deposition in brain and bone", *Magn Reson Imaging*, vol. 34, no. 10, pp. 1359–1365, 2016.

[99] J. Nocedal & S. J. Wright, *Numerical optimization*, springer ed., New York, 2006.

[100] A. R. Padhani, G. Liu, D. Mu-koh, T. L. Chenevert, H. C. Thoeny, B. D. Ross, M. V. Cauteren, D. Collins, D. A. Hammoud, G. J. S. Rustin, B. Taouli, "Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations", *Neoplasia*, vol. 11, no. 2, pp. 102–125, 2009.

[101] V. M. Pai, S. Rapacchi, P. Kellman, P. Croisille, H. Wen, "PCATMIP: enhancing signal intensity in diffusion-weighted magnetic resonance imaging", *Magn Reson Med*, vol. 65, pp. 1611–1619, 2011.

[102] S. C. Partridge, R. S. Murthy, A. Ziadloo, S. W. White, K. H. Allison, C. D. Lehman, "Diffusion tensor magnetic resonance imaging of the normal breast", *Magn Reson Imaging*, vol. 28, no. 3, pp. 320–328, 2010.

[103] S. C. Partridge, A. Ziadloo, R. Murthy, S. W. White, S. Peacock, P. R. Eby, W. B. DeMartini, C. D. Lehman, "Diffusion tensor MRI: preliminary anisotropy measures and mapping of breast tumors", *J Magn Reson Imaging*, vol. 31, no. 2, pp. 339–347, 2010.

[104] M. Paulides, P. Stauffer, E. Neufeld, P. Maccarini, A. Kyriakou, R. Canters, C. Diederich, J. Bakker, G. Van Rhoon, "Simulation techniques in hyperthermia treatment planning", *Int J Hyperther*, vol. 29, no. 4, pp. 346–357, 2013.

[105] J. Pluim, J. Maintz, M. Viergever, "Mutual information based registration of medical images: a survey", *IEEE T Med Imaging*, vol. 22, no. 8, pp. 1–21, 2003.

[106] M. Polfliet, W. Huizinga, S. Klein, J. D. Mey, "Pythagorean mean images for efficient groupwise registration", in *MICCAI*, pp. 26–33 (ICART workshop), 2015.

[107] D. H. J. Poot & S. Klein, "Detecting statistically significant differences in quantitative MRI experiments, applied to diffusion tensor imaging", *IEEE T Med Imaging*, vol. 34, no. 5, pp. 1164–1176, 2015.

[108] J. Qi, N. J. Olsen, R. R. Price, J. A. Winston, J. H. Park, "Diffusion-weighted imaging of inflammatory myopathies: polymyositis and dermatomyositis", *J Magn Reson Im*, vol. 27, no. 1, pp. 212–217, 2008.

[109] H. Ragheb, N. A. Thacker, J.-M. Guyader, S. Klein, N. M. DeSouza, A. Jackson, "The accuracy of ADC measurements in liver is improved by a tailored and computationally efficient local-rigid registration algorithm", *PLOS One*, vol. 10, no. 7, pp. 1–18 (e0132554), 2015.

[110] J. M. Ream, J. R. Dillman, J. Adler, S. Khalatbari, J. B. McHugh, P. J. Strouse, M. Dhanani, B. Shpeen, M. M. Al-Hawary, "MRI diffusion-weighted imaging (DWI) in pediatric small bowel Crohn disease: Correlation with MRI findings of active bowel wall inflammation", *Pediatr Radiol*, vol. 43, no. 9, pp. 1077–1085, 2013.

[111] Y. D. Reijmer, A. Leemans, S. M. Heringa, I. Wielaard, B. Jeurissen, H. L. Koek, G. J. Biessels, "Improved sensitivity to cerebral white matter abnormalities in Alzheimer's disease with spherical deconvolution based tractography", *PLoS One*, vol. 7, no. 8, pp. 1–8 (e44074), 2012.

[112] H. Robbins & S. Monro, "A stochastic approximation method", *Ann Math Stat*, vol. 22, no. 3, pp. 400–407, 1951.

[113] G. K. Rohde, S. Pajevic, C. Pierpaoli, "Multi-channel registration of diffusion tensor images using directional information", in *ISBI*, pp. 712–715, 2004.

[114] G. K. Rohde, S. Pajevic, C. Pierpaoli, P. Basser, "A comprehensive approach for multi-channel image registration", in *WBIR*, pp. 214–223, 2003.

[115] T. Rohlfing, C. R. Maurer, D. A. Bluemke, M. A. Jacobs, "Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint", *IEEE T Med Imaging*, vol. 22, no. 6, pp. 730–741, 2003.

[116] M. Rosenfeld, J. Emerson, J. Williams-Warren, M. Pepe, A. Smith, A. B. Montgomery, B. Ramsey, "Defining a pulmonary exacerbation in cystic fibrosis", *J Pediatr*, vol. 139, no. 3, pp. 359–365, 2001.

[117] D. Rueckert & J. A. Schabel, *Medical image registration*, Springer-Verlag, Berlin, 2011.

[118] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images", *IEEE T Med Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[119] D. Russakoff, C. Tomasi, T. Rohlfing, C. Maurer, T. Pajdla, J. Matas, "Image similarity using mutual information of regions", in *ECCV*, pp. 596–607, 2004.

[120] D. Seghers, E. D. Agostino, F. Maes, D. Vandermeulen, "Construction of a brain template from MR images using state-of-the-art registration and segmentation techniques", in *MICCAI*, Rennes, Brittany, pp. 696–703, 2004.

[121] C. E. Shannon, "A mathematical theory of communication", *Bell Syst Tech J*, vol. 27, pp. 379–423, 1948.

[122] J. Sijbers, A. J. D. Dekker, P. Scheunders, D. V. Dyck, "Maximum likelihood estimation of Rician distribution parameters", *IEEE T Med Imaging*, vol. 17, no. 3, pp. 357–361, 1998.

[123] J. G. Sled, A. P. Zijdenbos, A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data", *IEEE T Med Imaging*, vol. 17, no. 1, pp. 87–97, 1998.

[124] A. Sotiras, C. Davatzikos, N. Paragios, A. Sotiras, C. Davatzikos, N. Paragios, D. Medical, I. Registration, A. Sotiras, C. Davatzikos, S. Member, N. Paragios, "Deformable medical image registration: a survey", *IEEE T Med Imaging*, vol. 32, no. 7, pp. 1–39, 2013.

[125] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation", *IEEE T Automat Contr*, vol. 37, no. 3, pp. 332–341, 1992.

[126] Z. Špiclin, B. Likar, F. Pernuš, "Groupwise registration of multimodal images by an efficient joint entropy minimization scheme", *IEEE T Med Imaging*, vol. 21, no. 5, pp. 2546–2558, 2012.

[127] M. Staring, S. Klein, J. P. Pluim, "A rigidity penalty term for nonrigid registration", *Med Phys*, vol. 34, no. 11, pp. 4098–4108, 2007.

[128] M. Staring, U. A. van der Heide, S. Klein, M. A. Viergever, J. P. W. Pluim, "Registration of cervical MRI using multifeature mutual information.", *IEEE T Med Imaging*, vol. 28, no. 9, pp. 1412–1421, 2009.

[129] R. O. Suarez, O. Commowick, S. P. Prabhu, S. K. Warfield, "Automated delineation of white matter fiber tracts with a multiple region-of-interest approach", *NeuroImage*, vol. 59, no. 4, pp. 3690–3700, 2012.

[130] R. Tanaka, H. Horikoshi, Y. Nakazato, E. Seki, K. Minato, M. Iijima, M. Kojima, T. Goya, "Magnetic resonance imaging in peripheral lung adenocarcinoma: correlation with histopathologic features", *J Thorac Imag*, vol. 24, no. 1, pp. 4–9, 2009.

[131] B. Taouli & D.-M. Koh, "Diffusion-weighted MR imaging of the liver", *Radiology*, vol. 254, no. 1, pp. 47–66, 2010.

[132] B. Taouli, A. Sandberg, A. Stemmer, T. Parikh, S. Wong, J. Xu, V. S. Lee, "Diffusion-weighted imaging of the liver: comparison of navigator triggered and breathhold acquisitions", *J Magn Reson Imaging*, vol. 568, no. 30, pp. 561–568, 2009.

[133] P. Thévenaz & M. Unser, "Optimization of mutual information for multiresolution image registration", *IEEE T Med Imaging*, vol. 9, no. 12, pp. 2083–2099, 2000.

[134] D. Tibussek, C. Rademacher, J. Caspers, B. Turowski, J. Schaper, G. Antoch, D. Klee, "Gadolinium brain deposition after macrocyclic gadolinium administration: a pediatric case-control study", *Radiology*, vol. 285, no. 1, pp. 223–230, 2017.

[135] H. A. W. M. Tiddens, S. H. Donaldson, M. Rosenfeld, P. D. Paré, "Cystic fibrosis lung disease starts in the small airways: can we treat it more effectively?", *Pediatr Pulm*, vol. 45, no. 2, pp. 107–117, 2010.

[136] H. A. W. M. Tiddens, S. M. Stick, J. M. Wild, P. Ciet, G. J. M. Parker, A. Koch, J. Vogel-Claussen, "Respiratory tract exacerbations revisited: ventilation, inflammation, perfusion, and structure (VIPS) monitoring to redefine treatment", *Pediatr Pulm*, vol. 50, pp. S57–S65, 2015.

[137] N. H. Timm, *Applied multivariate analysis*, Springer, New-York, 2002.

[138] N. Timme, W. Alford, B. Flecker, J. M. Beggs, "Synergy, redundancy, and multivariate information measures: an experimentalist's perspective", *J Comput Neurosci*, vol. 36, pp. 119–140, 2014.

[139] A. Uitterdijk, T. Springeling, M. van Kranenburg, R. W. B. van Duin, I. Krabbendam-Peters, C. Gorsse-Bakker, S. Sneep, R. van Haeren, R. Verrijk, R.-J. M. van Geuns, W. J. van der Giessen, T. Markkula, D. J. Duncker, H. M. M. van Beusekom, "VEGF165A microsphere therapy for myocardial infarction suppresses acute cytokine release and increases capillary density, but does not improve cardiac function", *Am J Physiol-Heart C*, vol. 309, no. 3, pp. 396–406, 2015.

[140] T. Uto, Y. Takehara, Y. Nakamura, T. Naito, D. Hashimoto, N. Inui, T. Suda, H. Nakamura, K. Chida, "Higher sensitivity and specificity for diffusion-weighted imaging of malignant lung lesions without apparent diffusion coefficient quantification", *Radiology*, vol. 252, no. 1, pp. 247–54, 2009.

[141] N. van der Aa, A. Leemans, F. J. Northington, H. L. Van Straaten, I. C. Van Haastert, F. Groenendaal, M. J. N. L. Benders, L. S. De Vries, "Does diffusion tensor imaging-based tractography at 3 months of age contribute to the prediction of motor outcome after perinatal arterial ischemic stroke?", *Stroke*, vol. 42, no. 12, pp. 3410–3414, 2011.

[142] N. van der Aa, H. G. Ter Morsche, R. R. M. Mattheij, "Computation of eigenvalue and eigenvector derivatives for a general complex-valued eigensystem", *Electron J Linear Al*, vol. 16, pp. 300–314, 2007.

[143] J. Venn, "On the diagrammatic and mechanical representation of propositions and reasonings", *Philos Mag Series 5*, vol. 9, no. 59, pp. 1–18, 1880.

[144] T. Vercauteren, X. Pennec, A. Perchant, N. Ayache, "Symmetric log-domain diffeomorphic registration: a demons-based approach", in *MICCAI*, pp. 754–761, 2008.

[145] T. Vercauteren, X. Pennec, A. Perchant, N. Ayache, "NeuroImage diffeomorphic demons: efficient non-parametric image registration", *NeuroImage*, vol. 45, no. 1, pp. S61–S72, 2009.

[146] R. F. Verhaart, V. Fortunati, G. M. Verduijn, A. van der Lugt, T. van Walsum, J. F. Veenland, M. M. Paulides, "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", *Med Phys*, vol. 41, no. 12, pp. 1–11 (123302), 2014.

[147] R. F. Verhaart, V. Fortunati, G. M. Verduijn, T. van Walsum, J. F. Veenland, M. M. Paulides, "CT-based patient modeling for head and neck hyperthermia treatment planning: manual versus automatic normal-tissue-segmentation", *Radiother Oncol*, vol. 111, no. 1, pp. 158–163, 2014.

[148] M. A. Viergever, J. B. Maintz, S. Klein, K. Murphy, M. Staring, J. P. Pluim, "A survey of medical image registration - under review", *Med Image Anal*, vol. 33, pp. 140–144, 2016.

[149] P. Viola & W. Wells, "Alignment by maximization of mutual information", *Int J Comput Vision*, vol. 24, no. 2, pp. 137–154, 1997.

[150] C. Wachinger & N. Navab, "Simultaneous registration of multiple images: similarity metrics and efficient optimization", *IEEE T Pattern Anal*, vol. 35, no. 5, pp. 1221–1233, 2013.

[151] D. A. Walsh & C. I. Pearson, "Angiogenesis in the pathogenesis of inflammatory joint and lung diseases", *Arthritis Res*, vol. 3, no. 3, pp. 147–153, 2001.

[152] H.-C. Wang, J.-L. Hsu, A. Leemans, "Diffusion tensor imaging of vascular parkinsonism: structural changes in cerebral white matter and the association with clinical severity", *Arch Neurol*, vol. 69, no. 10, pp. 1340–1348, 2012.

[153] S. Watanabe, "Information theoretical analysis of multivariate correlation", *IBM J Res Dev*, vol. 4, no. 1, pp. 66–82, 1960.

[154] J. C. Waterton & L. Pylkkanen, "Qualification of imaging biomarkers for oncology drug development", *Eur J Cancer*, vol. 48, no. 4, pp. 409–415, mar 2012.

[155] J. West, J. Fitwpatrick, M. Wang, B. Dawant, "Comparison and evaluation of retrospective intermodality brain image registration techniques", *J Comput Assist Tomo*, vol. 21, no. 4, pp. 554–566, 1997.

[156] M. O. Wielpütz, M. Eichinger, M. Puderbach, "Magnetic resonance imaging of cystic fibrosis lung disease", *J Thorac Imag*, vol. 28, no. 3, pp. 151–159, 2013.

[157] M. O. Wielpütz, M. Puderbach, A. Kopp-Schneider, M. Stahl, E. Fritzsching, O. Sommerburg, S. Ley, M. Sumkauskaite, J. Biederer, H. U. Kauczor, M. Eichinger, M. A. Mall, "Magnetic resonance imaging detects changes in structure and perfusion, and response to therapy in early cystic fibrosis lung disease", *Am J Respir Crit Care Med*, vol. 189, no. 8, pp. 956–965, 2014.

[158] C. Yan, J. Xu, W. Xiong, Q. Wei, R. Feng, Y. Wu, Q. Liu, C. Li, Q. Chan, Y. Xu, "Use of intravoxel incoherent motion diffusion-weighted MR imaging for assessment of treatment response to invasive fungal infection in the lung", *Eur Radiol*, vol. 27, no. 1, pp. 212–221, 2017.

Summary

In this thesis, we develop medical image registration techniques for quantitative imaging and for multi-channel images.

In Chapter 2, we devise an image registration pipeline for the registration of diffusion-weighted MR (DW-MR) images acquired from the abdominal area of healthy volunteers with a free-breathing protocol. Spatial correspondence is not ensured in the acquired image datasets, mainly because of respiratory motion. The lack of spatial correspondence is apparent between the acquired images but also within each of them, due to the interleaved acquisition protocol chosen to limit the effect of cross-talk between slices. Our experiments, performed on five healthy volunteers with repeated acquisitions, show that image registration is an essential tool to ensure the extraction of reliable and reproducible apparent diffusion coefficient (ADC) quantitative measurements from the image datasets.

Chapter 3 focuses on DW-MR images acquired from 56 cystic fibrosis (CF) patients with respiratory tract exacerbation (RTE). The DW-MR images are acquired with a protocol close to the one used in Chapter 2, and are therefore registered using an image registration pipeline adapted from Chapter 2. Quantitative measurements are extracted based on the intravoxel incoherent motion (IVIM) theory, which contrary to the ADC technique, allows to quantify molecular diffusion and perfusion in a separate manner. We show that the IVIM quantitative parameter D^*, characterising blood perfusion, is a promising quantitative parameter for capturing treatment response in CF patients affected by RTE.

Chapter 4 makes it possible to apply one of the multivariate versions of mutual information in the context of groupwise image registration. Our interest in multivariate dissimilarity measures based on mutual information lies in the fact that pairwise mutual information is successfully used in common practice in a large number of applications, and is justified by the lack of groupwise equivalents or extensions used for image registration. This chapter shows that groupwise total correlation yields better registration accuracy results than pairwise registration based on mutual information, on six types of quantitative MR and dynamic imaging datasets containing between $G = 5$ and $G = 160$ images.

In Chapter 5, we develop a new groupwise scheme for the registration of datasets of multi-channel images, i.e. images for which multiple sub-images (called channels) are available. Contrary to the conventional pairwise multi-channel registration techniques, our groupwise multi-channel technique allows to register all channel images simultaneously, does not require the selection of a reference image space,

and assigns a single transformation to all channel images that belong together. Such channel images may be created by applying filtering, like we did through the use of the MIND feature images. Our groupwise multi-channel method also opens new promising perspectives for studies based on datasets of multi-channel images generated by the imaging modality itself.

This thesis proposes registration methods that improve the reliability and reproducibility of quantitative measurements extracted in a voxelwise manner from multiple medical images, and applies these registration methods to multiple medical imaging modalities focusing on different anatomical areas of interest or pathologies. In this thesis, we also design a novel registration scheme that allows the registration of images for which multiple channels exist. The image registration techniques that are developed are modular in the sense that the registration components can be tailored for a given application. Our work sheds a particular light on groupwise registration techniques, as these techniques might get an increasing interest in medical imaging in the coming years. In addition to bringing about advances in registration accuracy, groupwise registration could be an elegant and efficient way of decomplexifying image registration pipelines in a time marked by an inflation in the numbers and types of medical images acquired in the clinic.

Samenvatting

In dit proefschrift worden medische beeldregistratiemethoden ontwikkeld voor kwantitatieve beeldvoorming en meerkanaals beelden.

In hoofdstuk 2 wordt een beeldregistratiepijplijn bedacht voor de registratie van diffusie-gewogen MR (DW-MR) beelden, gevormd in de buik van gezonde vrijwilligers onder vrije ademhaling. De spatiale overeenstemming binnen de beeldgegevenssets is niet gegarandeerd vooral vanwege ademhalingsbeweging. Het gebrek aan spatiale overeenstemming is bemerkbaar tussen de verkregen beelden maar ook binnen elk beeld, ten gevolgen van de verweven beeldacquisitieprotocol gekozen om de effecten van overspraak tussen de 2D plakjes te beperken. Op experimenten gebaseerd op medische beelden van vijf gezonde vrijwilligers met herhaalde beeldacquisities blijkt dat beeldregistratie een onmisbaar instrument is om ervoor te zorgen dat de kwantitatieve apparente diffusie coëfficiënt (ADC) waarde berekend uit de DW-MR beelden betrouwbaar en reproduceerbaar zijn.

Hoofdstuk 3 richt zich op DW-MR beelden van 56 cystische fibrose (CF)-patiënten met verergeringen van het respiratoire kanaal (RTE). De DW-MR beelden worden gevormd met een protocol dat vergelijkbaar is met dat van hoofdstuk 2, en worden vervolgens geregistreerd met een beeldregistratiepijplijn aangepast van die van hoofdstuk 2. Kwantitatieve metingen worden berekend op basis van de theorie van intravoxel incoherente beweging (IVIM). In tegenstelling tot ADC methoden, IVIM maakt het mogelijk om de moleculaire diffusie en de perfusie in een aparte manier te kwantificeren. Wij tonen aan dat het IVIM kwantitatieve parameter D^*, die bloedperfusie kwantificeert, een veelbelovende kwantitatieve parameter is om de behandelrespons voor CF-patiënten met RTE te bepalen.

In hoofdstuk 4 wordt een methode beschreven, die het mogelijk maakt om een multivariate versie van mutuele informatie in de context van groepsgewijze registratie toe te passen. Onze belangstelling voor dissimilariteitsmaten afstammend uit mutuele informatie ligt in het feit dat paarsgewijze registratiemethoden gebaseerd op mutuele informatie doorgaans succesvol zijn voor een groot aantal toepassingen, en wordt gerechtvaardigd door het gebrek aan groepsgewijze equivalenten of uitbreidingen voor beeldregistratie. Dit hoofdstuk toont aan dat groepsgewijze totale correlatie een betere nauwkeurigheidsresultaten oplevert dan paarsgewijze registratiemethoden gebaseerd op mutuele informatie, op zes soorten kwantitatieve en dynamische beeldvorming gegevenssets bestaande uit tussen $G = 5$ en $G = 160$ beelden.

In hoofdstuk 5 wordt een nieuw groepsgewijs schema ontwikkeld voor de registratie van gegevenssets van meerkanaals beelden, d.w.z. beelden waarvoor meer-

dere subbeelden (die ook wel kanalen genoemd worden) beschikbaar zijn. In tegenstelling tot de conventionele paarsgewijze meerkanaals registratiemethoden, onze groepsgewijze meerkanaalstechniek maakt het mogelijk om alle kanalen gelijktijdig te registreren, zonder dat een referentiebeeld hoeft te worden gekozen, en één transformatie wordt toegewezen aan alle kanalen die bij elkaar behoren. Zulke kanaalbeelden kunnen onder andere samengesteld worden door filtering, zoals uitgevoerd in deze studie met MIND kenmerkbeelden. Onze groepsgewijze methode biedt ook veelbelovende perspectieven voor studies gebaseerd op gegevenssets van meerkanaals beelden gecreëerd door de beeldvormingstechniek zelf.

Dit proefschrift draagt bij aan de ontwikkeling van registratiemethoden die de betrouwbaarheid en reproduceerbaarheid verbeterd van kwantitatieve metingen verkregen op een voxel-wijze manier uit meerdere medische beelden. Deze registratiemethoden worden aan beelden van meervoudige beeldvormingstechnieken, anatomische gebieden of ziektes toegepast. Ook hebben wij een nieuw registratieschema ontworpen voor beelden gemaakt van meerdere kanalen. De registratiemethoden beschreven in dit proefschrift zijn modulair, in de zin dat de registratiecomponenten kunnen per toepassing aangepast worden. Ons werk hecht grote waarde aan groepsgewijze registratiemethoden mede gelet op het feit dat ze tot verhoogde belangstelling in de komende jaren zouden leiden. Naast de vooruitgang wat betreft registratienauwkeurigheid, groepsgewijze registratie zou een elegante en efficiënte manier om de beeldregistratie pijplijnen te decomplexificeren in een periode die in het teken staat van de groeiende aantallen en soorten medische beelden die in de kliniek gevormd worden.

PhD Portfolio

PhD period 2012-2018
Departments Medical Informatics | Radiology & Nuclear Medicine
Research school ASCI

In-depth courses

Front-end vision and multiscale image analysis – ASCI research school
TU Eindhoven November 11–15, 2013

Presentation course (4 sessions)
Erasmus MC, Rotterdam February–March 2013

International computer vision summer school (ICVSS)
Le Castella, Italy July 14–20, 2013

Advanced pattern recognition – ASCI research school
TU Delft April 7–11, 2014

Scientific writing course (10 sessions)
Erasmus MC, Rotterdam November 2014–February 2015

Multimodal imaging towards individualized radiotherapy treatments
TU Delft July 8–10, 2014

International conferences

European Society for Magnetic Resonance in Medicine and Biology (ESMRMB) – Poster
Toulouse, France October 3–5, 2013

SPIE Medical Imaging: Image Processing – Oral presentation
San Diego, USA February 15–20, 2014

IEEE International Symposium on Biomedical Imaging (ISBI) – Oral presentation
New York, USA April 16–19, 2015

7th International Workshop on Biomedical Image Registration (WBIR) – Oral presentation
Las Vegas, USA July 1, 2016

Research seminars

Biomedical Imaging Group Rotterdam (BIGR) seminars – Bi-weekly (3 oral presentations)
Erasmus MC, Rotterdam 2012–2016

Medical informatics research lunches – Bi-weekly (2 oral presentations)
Erasmus MC, Rotterdam 2012–2016

QuIC-ConCePT consortium presentations
Brussels, Belgium 2012
London, UK 2013
Manchester, UK 2014

Oncology meetings, BIGR group – Monthly (4 oral presentations)
Erasmus MC, Rotterdam, 2013–2014

Joint BIGR-LUMC image registration (RAMBO) meetings – Bi-weekly (5 oral presentations)
Erasmus MC, Rotterdam and LUMC, Leiden 2013–2016

Medical imaging symposium for PhD students (MISP)
LUMC, Leiden May 8, 2014

Interdisciplinary seminar on Medical Imaging and Astronomy – Event organized by researchers of BIGR and Leiden University
Leiden November 11, 2015

Teaching experience

Medical image processing practical works for medical students
Erasmus MC, Rotterdam 2013

Supervision of Jian Yang (ENSEEIHT, Toulouse, France) on a Master's project entitled: "Application of groupwise PCA-based dissimilarity measures on multimodal datasets of medical images"
Erasmus MC, Rotterdam June–September, 2014

Others

Referee activities for the following international scientific journals: IEEE Transactions on Medical Imaging, and Medical Image Analysis.

Publications

Publications in international journals

J.-M. Guyader, L. Bernardin, N. H. M. Douglas, D. H. J. Poot, W. J. Niessen and S. Klein, "Influence of image registration of ADC images computed from free-breathing diffusion MRIs of the abdomen", *Journal of Magnetic Resonance Imaging*, vol. 42, no. 2, pp. 315–330, 2015

H. Ragheb, N. Thacker, **J.-M. Guyader**, S. Klein, N. deSouza and A. Jackson, "The accuracy of ADC measurements in liver is improved by a tailored and computationally efficient local-rigid registration algorithm", *PLOS One*, vol. 10, no. 7, pp. e0132554, 2015

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

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

J.-M. Guyader, W. Huizinga, V. Fortunati, D. H. J. Poot, J.F. Veenland, M. M. Paulides, W. J. Niessen and S. Klein, "Groupwise multi-channel image registration", *IEEE Journal of Biomedical and Health Informatics*, *in press*

J.-M. Guyader, P. Ciet, A. Mazzaro, T. Feiweier, G. Morana and S. Klein, "Monitoring of cystic fibrosis lung disease based on intravoxel incoherent motion and image registration", *submitted*

Publications in international conference proceedings

J.-M. Guyader, L. Bernardin, N. H. M. Douglas, D. H. J. Poot, W.J. Niessen and S. Klein, "Influence of image registration on ADC images computed from free-

breathing diffusion MRIs of the abdomen", in *SPIE Medical Imaging*, vol. 9034, pp. 90340O-1–90340O-6, 2014

W. Huizinga, D. H. J. Poot, **J.-M. Guyader**, H. Smit, M. van Kranenburg, R. J. van Geuns, A. Uitterdijk, H. van Beusekom, B. F. Coolen, A. Leemans, W. J. Niessen and S. Klein, "Non-rigid groupwise image registration for motion compensation in quantitative MRI", in *WBIR, International Workshop on Biomedical Image Registration*, pp. 184–193, 2014

J.-M. Guyader, W. Huizinga, V. Fortunati, J. F. Veenland, M. Paulides, W. J. Niessen, S. Klein, "Groupwise image registration of multimodal head-and-neck images", in *ISBI, IEEE International Symposium on Biomedical Imaging*, pp. 730–733, 2015

J.-M. Guyader, W. Huizinga, V. Fortunati, D. H. J. Poot, M. van Kranenburg, J. F. Veenland, M. Paulides, W. J. Niessen, S. Klein, "Total correlation-based groupwise image registration for quantitative MRI", in *WBIR, International Workshop on Biomedical Image Registration*, pp. 186–193, 2016

Conference abstracts

J.-M. Guyader, L. Bernardin, N. H. M. Douglas, D. H. J. Poot, W. J. Niessen and S. Klein, "Influence of image registration on the quality of computed apparent diffusion coefficients (ADC) in free-breathing diffusion MRI of the abdomen", in *ESMRMB, European Society for Magnetic Resonance in Medicine and Biology*, no. 162, p. 131, 2013

J.-M. Guyader, L. Bernardin, N. H. M. Douglas, D. H. J. Poot, W. J. Niessen and S. Klein, "Influence of image registration on the quality of computed apparent diffusion coefficients in free-breathing diffusion MRI of the abdomen", in *NVPHBV, Dutch Society for Pattern Recognition and Image Processing*, Spring meeting, 2013

H. Ragheb, N. A. Thacker, **J.-M. Guyader**, S. Klein, A. Jackson, "An efficient motion correction method for improved ADC estimates in the abdomen", in *ISMRM, International Society for Magnetic Resonance in Medicine*, no. 2984, 2015

J.-M. Guyader, P. Ciet, A. Mazzaro, H. A. W. M. Tiddens, G. Morana, S. Klein, "Can intravoxel incoherent motion (IVIM) be used to monitor treatment response in cystic fibrosis (CF) respiratory tract exacerbation (RTE)?", in *ECR, European Congress of Radiology*, 2018

Curriculum Vitae

Jean-Marie Guyader was born in Gwengamp/Guingamp, Brittany, France in 1989. He received his Engineer Diploma in computer science, electronics and biomedical technologies from the Institut supérieur de l'électronique et du numérique (ISEN) Brest, Brittany, France. He also holds a Master's degree in applied physics from the École supérieure de physique et de chimie industrielles de la ville de Paris (ESPCI), France.

Jean-Marie started his PhD trajectory in September 2012 within the Departments of Medical Informatics and Radiology of the Erasmus Medical Centre Rotterdam. The results of Jean-Marie's research are described in this thesis.