

DOCTORAL THESIS

Manifold Alignment for Imaging and Modelling Respiratory Motion

By Christian Frederik Baumgartner

King's College London
School of Medicine
Division of Imaging Sciences & Biomedical Engineering



1st Supervisor: Andrew P. King
King's College London
2nd Supervisor: Daniel Rueckert
Imperial College London
3rd Supervisor: Jamie R. McClelland
University College London

October 7, 2015

“The price of anything is the amount of life you exchange for it.”

— Henry Thoreau

“You have brains in your head. You have feet in your shoes. You can steer yourself any direction you choose.”

— Dr. Seuss

Abstract

Respiratory motion affects a wide range of techniques in the field of medical image acquisition and analysis. In image-guided interventions it may cause misalignment of static road maps with the patient’s anatomy. In imaging such as magnetic resonance (MR) imaging or positron emission tomography (PET) it may cause the images to appear blurred which may impede disease diagnosis and staging.

Manifold learning is a powerful tool for the non-linear dimensionality reduction of imaging data, which can be used to uncover the data’s dominating sources of motion. By aligning low-dimensional embeddings of multiple datasets, which vary due to the same motion, in a joint low-dimensional space accurate correspondences between the datasets can be established.

In the first part of this thesis, manifold alignment is investigated for the robust reconstruction of high-resolution 4D (3D+time) MR imaging sequences of respiratory motion from sequentially acquired 2D MR slices. In particular, a novel groupwise manifold alignment scheme is presented which outperforms two current state-of-the-art reconstruction techniques. From such 4D MR images very accurate motion estimates are derived, which, in turn, are used to correct for motion in simulated PET-MRI data.

In the second part of this thesis, a patient-specific respiratory motion model presented based on groupwise manifold alignment. Such a motion model can be used to correct for 3D organ motion during an image-guided intervention where only 2D images are available. It is shown that the aligned low-dimensional representations obtained using manifold alignment may be viewed directly as a surrogate-driven motion model. By updating this low-dimensional manifold with points obtained from new 2D imaging data, the model can automatically adapt to previously unseen breathing patterns.

Lastly, in the third part of this thesis, a novel manifold alignment method is outlined which does not require any of the datasets to be comparable in the image space and can thus be used to align the manifolds of visually distinct datasets. Using this method the 4D MR reconstruction from multiple 2D MR slices is revisited. Furthermore, the methodology is applied to time-resolved compounding of multiple 3D ultrasound views of the same organ.

Acknowledgements

Over the past years I have received support and encouragement from a great number of individuals. Firstly, I would like to express my deep gratitude to my supervisor, Dr Andrew King, who supported me in all aspects of my project. It is hard to imagine how I could have managed without his guidance. I am also very thankful to my secondary supervisors, Professor Daniel Rueckert and Dr Jamie McClelland for our frequent meetings in which they gave me the benefit of their vast experience, and for all the time they invested in proof reading drafts. In particular, I would like to thank Daniel for giving me the opportunity to work as a research associate at Imperial College upon finishing my PhD.

My sincere thanks go to my colleagues in the Division of Imaging Sciences & Biomedical Engineering who helped me not only by advising on technical issues, but also made me feel at home at St. Thomas' Hospital. In particular, I am grateful to Andrew Aitken and Christoph Kolbitsch for teaching me how to operate the MR scanner, setting up protocols for me, and helping me with all MR related (and unrelated) questions and problems I encountered along the way. My thanks also go to Devis Peressutti for numerous encouraging discussions and for letting me steal all his templates. I would also like to thank Markus, Alberto, Jack, Gastao, Ghislain, Jedrek and Nick for all the Lower Marsh lunches and after work pints.

I am greatly indebted to my parents who have believed in me in a time when academic success seemed unlikely and I was on the brink of stopping to believe in myself. Without their continuing love and support I could not have achieved anything.

Lastly, I would like to thank Lisa who, to my great delight and relief, decided to move to London with me three years ago. She supported me during this time not only as a colleague through countless discussions and with her incredible insights, but more importantly, as a friend and partner by sharing the good and the bad of this journey.

Christian Baumgartner, October 2015

Contents

1	Introduction	9
1.1	Motivation	9
1.2	Contributions	10
1.3	Outline	10
2	Challenges of Respiratory Motion	13
2.1	Physiology of Respiratory Motion	13
2.2	Extent and Directions of Respiratory Motion	14
2.3	Respiratory Variabilities	15
2.3.1	Inter-cycle Variability	16
2.3.2	Intra-cycle Variability	17
2.4	Breath-holding	17
2.5	Problems Caused By Respiratory Motion	18
2.5.1	Interventions	19
2.5.2	Imaging of Anatomy with Respiratory Motion	20
2.6	Discussion	21
3	Literature Review	22
3.1	Correcting for Respiratory Motion	23
3.1.1	Breath-Holding	23
3.1.2	External Gating Techniques	23
3.1.3	Self-Gating Methods	26
3.1.4	Motion Correction Using Another Modality	27
3.2	Imaging Respiratory Motion	30
3.2.1	Dynamic Imaging	30
3.2.2	Slice-Stacking	31
3.3	Patient-Specific Motion Modelling	31
3.3.1	Model Calibration	32
3.3.2	Model Formation	33

3.3.3	Model Application	37
3.4	Manifold Learning	39
3.4.1	Introduction	39
3.4.2	Manifold Learning For Understanding Physiological Motion	40
3.5	Manifold Alignment	41
3.5.1	Methods for Alignment	43
3.5.2	Types of Supervision	44
3.6	Discussion	45
4	Background and Theory	48
4.1	Manifold Learning on One Dataset	49
4.2	The Problem of Misaligned Embeddings	51
4.3	Manifold Alignment of Multiple Datasets	52
4.3.1	Two-Step Alignment	52
4.3.2	One-Step Alignment	53
5	Simultaneous Groupwise Manifold Alignment	56
5.1	Introduction	56
5.2	Materials and Methods	58
5.2.1	Slice-by-Slice Acquisition	58
5.2.2	Manifold Alignment for Multiple Datasets	59
5.2.3	Choices for Inter-Dataset Kernel	62
5.2.4	4D Volume Reconstruction	67
5.3	Experiments	68
5.3.1	Compared Methods	68
5.3.2	Comparison of Manifold Alignment Methods	70
5.3.3	Comparison to State-of-the-Art Slice-Stacking Methods	72
5.4	Results	73
5.4.1	Comparison of Manifold Alignment Methods	73
5.4.2	Comparison to State-of-the-Art Slice-Stacking Methods	77
5.5	Discussion	79
6	Autoadaptive Motion Modelling	90
6.1	Introduction	90
6.1.1	Motion Modelling	91
6.2	Materials and Methods	93
6.2.1	Calibration Scan	94
6.2.2	Motion Model Formation	95

6.2.3	Model Updating and Adaptivity	98
6.3	Experiments and Results	101
6.3.1	Parameter Choices	102
6.3.2	Experiment 1: Synthetic Training Adaptivity	103
6.3.3	Experiment 2: Synthetic Adaptivity to New Breathing Pattern	106
6.3.4	Experiment 3: Adaptivity on Real Data	109
6.4	Discussion	112
7	Self-Aligning Manifolds	115
7.1	Introduction	115
7.2	Method	117
7.2.1	Similarity Without Correspondences	117
7.2.2	Simultaneous Embedding of Many Datasets	119
7.3	Applications	120
7.3.1	Application to 4D magnetic resonance (MR) Reconstruction	120
7.3.2	Application to 3D Ultrasound Gating and Compounding	121
7.4	Experiments and Results	121
7.4.1	Experiment 1: Synthetic 4D MR Reconstruction	121
7.4.2	Experiment 2: 4D MR Reconstruction from Real Data	122
7.4.3	Experiment 3: 3D Ultrasound Compounding	123
7.5	Discussion	123
8	Conclusions	129
8.1	Summary of Novel Contributions	129
8.2	Current Limitations and Future Directions	130
8.3	Final Remarks	132
A	Additional Proofs	134
A.1	Extension of LLE to Two Datasets	134
B	Generation of Synthetic Data	137
B.1	Generation of Simple Synthetic MR Data (TYPE 1)	137
B.2	Generation of Synthetic PET Data	138
B.3	Generation of Motion Model-Based Synthetic MR Data (TYPE 2)	139
B.3.1	Synthetic 2D MR Slices	142
B.3.2	Synthetic Motion Fields	143
C	Summary of data used in this paper	144

List of Figures

2.1	Schematic of inspiration and expiration.	14
2.2	Example of respiratory traces with variability.	16
2.3	Lung schematic with 21 tumour trajectories.	18
3.1	Comparison of 2D and 3D MR acquisitions of the lungs.	29
3.2	Schematic illustration of motion model formation and application.	32
3.3	Manifold learning on the ‘Swiss roll’ manifold.	40
4.1	Example illustrating the problem of misaligned manifold embeddings.	52
5.1	Illustration of 4D MR reconstruction using manifold alignment.	58
5.2	Schematic illustration of groupwise manifold alignment.	62
5.3	Transformations involved in the computation of the registration based similarity kernel.	64
5.4	Sparsification of similarity kernel by one-to-one node matching.	67
5.5	Results of the tuning of parameter μ	74
5.6	Average error over all volunteers for all methods evaluated in the experiments on synthetic data.	74
5.7	Visual results of synthetic evaluation of the manifold alignment techniques.	75
5.8	Visual results of evaluation of the manifold alignment techniques on real data.	76
5.9	Visual results of synthetic evaluation of groupwise manifold learning and two state-of-the-art techniques.	78
5.10	Visual results for evaluation of groupwise manifold learning and two state-of-the-art techniques on real data.	87
5.11	Examples of aligned manifolds and manifolds aligned using the SGA.REG technique.	88
5.12	Examples of motion corrected synthetic positron emission tomography (PET) volumes.	88
5.13	SUV line profiles through the tumour.	89

6.1	Schematic of traditional, and the novel autoadaptive, motion model paradigm.	92
6.2	Schematic of group connections through simultaneous manifold embeddings.	94
6.3	Derivation of similarity kernel based on motion in slice overlap.	97
6.4	Schematic of the connection of slice positions by means of pairwise embedding and propagation of respiratory information through the manifolds.	99
6.5	Schematic illustrating how the κ nearest are found in the manifold embedding.	101
6.6	Error curves of synthetic validation on normal breathing data.	105
6.7	Error curves of synthetic validation on new breathing pattern.	108
6.8	Error curves of validation on real data.	110
6.9	Example of respiratory drift compensated for by the AAMM technique. . .	111
7.1	Results of synthetic 4D MR reconstruction experiment.	125
7.2	Example of joint embedding of all sagittal slice positions.	126
7.3	Results of the real 4D MR reconstruction.	127
7.4	Results of ultrasound (US) experiment.	128
B.1	Coronal slices at the same A-P position through components of the synthetic PET simulation.	140
B.2	Overview of synthetic slice-by-slice data generation of TYPE 2.	141

List of Tables

3.1	Overview of reviewed motion models.	35
3.2	Ability to capture respiratory variabilities and adaptivity of reviewed models.	36
3.3	Overview of reviewed manifold alignment techniques.	42
4.1	List of frequently used mathematical notations.	49
5.1	Summary of experiments in Chapter 5.	68
5.2	Quantitative results of synthetic evaluation of the manifold alignment techniques.	84
5.3	Quantitative results of evaluation of the manifold alignment techniques on real data.	85
5.4	Quantitative results of synthetic evaluation of groupwise manifold learning and two state-of-the-art techniques.	86
5.5	Quantitative results for evaluation of groupwise manifold learning and two state-of-the-art techniques on real data.	86
6.1	Error figures of synthetic validation.	104
6.2	Error figures of synthetic validation on deep breathing data.	107
6.3	Error figures of validation on real data.	110
C.1	Summary of all the data used in this thesis.	145

Acronyms

A-P	<i>Anterior-Posterior</i>
AAMM	<i>Autoadaptive Motion Model</i>
CCA	<i>Canonical Correlation Analysis</i>
CT	<i>Computed (X-ray) Tomography</i>
EPI	<i>Echo Planar Imaging</i>
FDG	<i>Fluorodeoxyglucose</i>
GRE	<i>Gradient Echo</i>
GT	<i>Ground-Truth</i>
HIFU	<i>High-Intensity Focused Ultrasound</i>
HLLE	<i>Hessian Locally Linear Embedding</i>
ICP	<i>Iterative Closest Point</i>
LEM	<i>Laplacian Eigenmaps</i>
LLE	<i>Locally Linear Embedding</i>
LLSR	<i>Linear Least Squares Regression</i>
LOO	<i>Leave-One-Out</i>
MCIR	<i>Motion Corrected Image-Reconstruction</i>
MIP	<i>Maximum Intensity Projection</i>
MR	<i>Magnetic Resonance</i>
NCC	<i>Normalised Cross Correlation</i>

OSEM	<i>Ordered Subsets Expectation Maximisation</i>
PCA	<i>Principal Component Analysis</i>
PET	<i>Positron Emission Tomography</i>
PLSR	<i>Partial Least Squares Regression</i>
R-L	<i>Right-Left</i>
RPM	<i>Varian Real Time Position Management System</i>
RT	<i>Radiotherapy</i>
RTA	<i>Reconstruct Transform-Average</i>
S-I	<i>Superior-Inferior</i>
SAM	<i>Self-Aligning Manifolds</i>
SGA	<i>Simultaneous Groupwise Manifold Alignment</i>
SUV	<i>Standardised Uptake Value</i>
US	<i>Ultrasound</i>

Chapter 1

Introduction

1.1 Motivation

Respiratory motion is one of the most significant sources of thoracic and abdominal organ motion and can significantly impact on medical treatments and imaging techniques. In imaging techniques such as MR imaging or PET respiration can negatively affect image quality causing blurring of structures or tumours which causes difficulties in disease detection, diagnosis and staging. In interventions such as radiotherapy or cardiac catheterisations which require, or are aided by image-based planning prior to the treatment, motion induced by respiration may cause the target area to deviate from the treatment plan. This may reduce the accuracy of the treatment and potentially cause harm to the patient.

Although approximately cyclic, breathing motion is not perfectly repeatable from breathing cycle to breathing cycle. In fact significant variations exist between breathing cycles, and also between inhalation and exhalation, impeding the accuracy of commonly used strategies to account for respiratory motion.

Some imaging modalities can be dynamically acquired in close to real-time allowing imaging of patients under free-breathing. For example, 2D MR slices can be acquired using standard sequences in less than 200 ms. In order to image volumes rather than planes, slice stacking approaches aim to combine multiple slices from different imaging planes into coherent 3D images. Due to the respiratory motion variabilities this is a challenging task. A closely related challenge is the compounding of free breathing 3D US data from different views.

In image-guided treatments it has been proposed to use *patient-specific motion models* to account for respiratory motion. Such models are calibrated using image data before the treatment and can then be used during the treatment to obtain close to real-time motion estimates by acquiring a simpler respiratory surrogate signal. Although many such motion models can account for some of the respiratory variations, most are restricted to the

breathing patterns observed before the treatment and may lose validity as the treatment goes on due to changing breathing patterns.

Dimensionality reduction techniques offer a powerful tool to analyse respiratory motion. In particular, non-linear dimensionality reduction techniques, collectively known as *manifold learning*, have been successfully applied to this task. *Manifold alignment* techniques are an extension of manifold learning methods which allow reduction of the dimensionality of multiple datasets into a single consistent low-dimensional space. Manifold learning and alignment techniques provide a possible tool to estimate respiratory motion directly from medical imaging data.

1.2 Contributions

The foremost contribution of this thesis is to investigate the potential of manifold alignment techniques to account for respiratory motion in various applications in medical image analysis.

In particular, several novel methods are proposed to establish correspondences between large numbers of medical imaging datasets in the low-dimensional space. Such correspondences are used for the following applications:

- To reconstruct accurate 4D MR sequences from sequentially acquired 2D MR slice data from different imaging planes. In contrast to dynamic 3D MR images, such 2D slices have particularly advantageous properties in the lungs, where they can visualise blood flow through the vasculature. When present in the reconstructed 3D images, these features allow for a very accurate estimation of the motion inside the whole thorax including the lungs.
- To accurately correct PET data for respiratory motion using simultaneously acquired 2D MR data in a hybrid PET/MR scenario.
- To simultaneously gate and compound multiple 3D ultrasound sequences of the same organ from different views.
- To implement a novel type of motion model which has the inherent ability to continually adapt to changing breathing patterns during image-guided treatments, allowing for increased accuracy in long treatment sessions.

1.3 Outline

The thesis is organised the eight chapters, which are outlined in detail below.

Chapter 2 provides a characterisation of respiratory motion and its variabilities and examines its impact on various clinical applications in medical imaging and treatments.

Chapter 3 contains a literature review which focuses on existing solutions to the problem of respiratory motion, including simple techniques to correct for respiratory motion, methods to image respiratory motion, and patient-specific motion modelling techniques. Furthermore, existing manifold learning and manifold alignment techniques and their applications in medical imaging are discussed.

Chapter 4 introduces the theory of two basic approaches to manifold alignment of two datasets and establishes the notation used throughout the remainder of this thesis.

Chapter 5 examines a number of extensions of basic manifold alignment methods to multiple datasets for the reconstruction of 4D MR volumes from coronal 2D slice-by-slice data. In particular, a novel groupwise manifold alignment approach is proposed which allows for robust and accurate reconstruction of such data. It is furthermore demonstrated how the approach can be used to correct PET data using simultaneously acquired MR data in a hybrid PET/MR scenario.

Chapter 6 describes how manifold alignment methodology can be used to implement an *autoadaptive* motion model which has the ability to automatically adapt to changing breathing patterns in lengthy image-guided treatments.

Chapter 7 describes a novel manifold alignment method which can be used to establish correspondences between dynamic imaging data from different views or different parts of the anatomy, which are too dissimilar to be compared in image-space. The technique is demonstrated for 4D MR reconstruction from sequentially acquired *sagittal* 2D MR slices, which is a more challenging problem than for coronal slices, and for the simultaneous gating and compounding of ultrasound images acquired from two different views.

Chapter 8 summarises the main contributions presented in this thesis and describes some limitations of the proposed methods. Moreover, a discussion on the clinical impact of the methods is provided along with an overview of future directions. Finally, conclusions are drawn.

In addition to the above chapters this thesis contains three appendices. Appendix A contains some additional proofs pertaining to Chapter 4. Appendix B describes the different approaches used to generate synthetic data for validating the approaches proposed in this

thesis. Lastly, Appendix C contains a summary of all the medical image datasets used in the experiments of Chapters 5-7.

Chapter 2

Challenges of Respiratory Motion

The main focus of this thesis is human respiratory motion and solutions to some of the problems it can cause in various aspects of medical image analysis. In Sections 2.1-2.4 the characteristics of this motion are more closely analysed before going into the details about the problems respiratory motion poses in medical imaging and image-guided treatments in Section 2.5.

2.1 Physiology of Respiratory Motion

Respiratory motion drives the intake and exchange of oxygen and carbon dioxide between an organism and the environment, which is required to keep the organism alive.

The passage of air into the lungs to supply the body with oxygen is known as inhalation or inspiration, and the passage of air out of the lungs to expel carbon dioxide is known as exhalation or expiration. In physiology this process is collectively called breathing or ventilation. The motion driving this process is usually referred to as respiratory motion.

Breathing is a mostly involuntary motion which is controlled by the autonomic nervous system, but can be voluntarily overridden to a certain degree, for example for short periods of breath-holding [Keall et al., 2006]. During inspiration, the diaphragmatic muscles in the abdomen contract, lowering the diaphragm. The space thus created is filled by the lung. Under certain circumstances the external intercostal muscles of the thorax help with inspiration by actively elevating the chest. During expiration, the diaphragm relaxes and is pushed up by the abdominal pressure. The muscles of the abdominal wall, and the intercostal muscles assist this movement [Faller and Schünke, 2004]. These two processes are illustrated in Fig. 2.1.

The contributions of the different muscle groups vary during inhalation and exhalation.

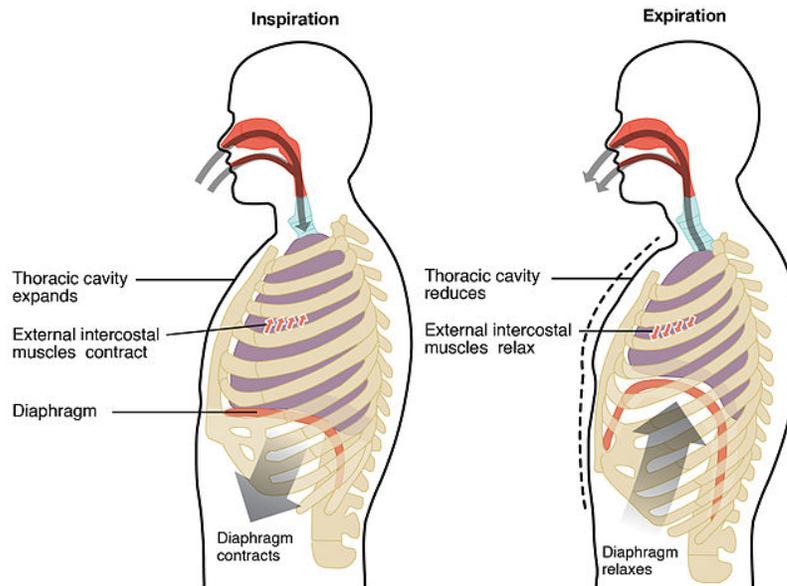


Figure 2.1: Sagittal schematic of inspiration and expiration indicating the movements of the diaphragm and external intercostal muscles¹.

tion, and with different breathing patterns. During normal, quiet breathing most subjects breathe predominantly with the diaphragm when supine and predominately with the thoracic muscles when upright [Sharp et al., 1975]. Fast and shallow breathing is accomplished by employing mostly the thoracic muscles, probably because the intercostal muscles are capable of more rapid action than the diaphragm and abdominal muscles [Sharp et al., 1975]. Normal breathing involves typical lung volume changes of around 20%. However, it has been found that the increase in lung volume can be between three to four times larger during deep breathing. Breathing patterns can and frequently do vary in magnitude, periodicity and regularity during a single imaging session [Keall et al., 2006]. Among other organs the lungs, esophagus, liver, pancreas, breasts, prostate and kidneys are known to move due to respiratory motion, thus affecting a wide range of different medical applications [McClelland et al., 2013]. The focus of this thesis lies mainly in imaging and accounting for respiratory motion of the thorax (i.e. primarily the lungs) and the liver, although the techniques developed here may also have applications in other organs.

2.2 Extent and Directions of Respiratory Motion

A number of studies have investigated the extent of respiratory motion. Many of these studies examined the motion of tumours or fiducial markers implanted close to tumours

¹This illustration is available from the Wikimedia Commons under the Creative Commons licence: https://commons.wikimedia.org/wiki/File:2316_Inspiration_and_Expiration.jpg

rather than the motion of the organs themselves, and the overall organ motion was assumed to correlate with that of the tumour. A good review can be found in Keall et al. [2006]. Generally it has been found that tumours higher up in the thorax (e.g. upper lung lobes) exhibit smaller motions than tumours in the lower lung lobes or the liver [Suh et al., 2008; Seppenwoolde et al., 2002; Plathow et al., 2004; Park et al., 2012]. Furthermore, on average respiratory motion is largest in the superior-inferior (S-I) direction, followed by the anterior-posterior (A-P) and then the right-left (R-L) directions [Suh et al., 2008; Seppenwoolde et al., 2002; Park et al., 2012; Korin et al., 1992] and there are typically no significant differences in respiratory motion between the left and right hemi-thoraxes [Plathow et al., 2004; Korin et al., 1992].

The actual peak-to-trough displacements reported vary from study to study, presumably due to the typically small number of datasets. The largest study of thoracic tumour motion to date was performed by Suh et al. [2008] and included multiple treatment fractions of 30 patients who underwent stereotactic radiotherapy. The authors reported an average peak-to-trough distance of 8 mm for the lower lung lobes and 3.6 mm for the upper lung lobes. Furthermore, it was found that the S-I motion was dominant in 57% of patients, the A-P motion in 27% of patients, and the R-L motion in 16% of patients. In a similar large study involving 20 patients, Seppenwoolde et al. [2002] reported tumour motion in the S-I direction of 12 mm on average in the lower lobes and 2 mm in the upper lobes. The authors found that the motion in the R-L and A-P directions was similar for upper and lower lobes and tended to be smaller than the S-I motion with average figures of 1.2 mm and 2.2 mm, respectively. Plathow et al. [2004] found that tissue mobility was significantly larger in healthy lung tissue compared to tissue affected by tumours. Therefore, the figures quoted above (for patients) may be larger for healthy subjects. Park et al. [2012] investigated the extent of tumour motion in the liver for 20 radiotherapy patients by tracking implanted fiducial markers and found average motion ranges of 16.5 mm, 5.3 mm and 2.8 mm in the S-I, A-P, and R-L directions, respectively. The authors found that the S-I and A-P motions were highly correlated whereas the R-L motion had a more variable relationship to the other directions.

2.3 Respiratory Variabilities

Breathing is an approximately periodic motion with typical cycle lengths of 3 to 5 seconds [Seppenwoolde et al., 2002]. However, respiratory motion can be characterised by significant variations between the inhale and exhale motion (known as hysteresis or intra-cycle variability) [Blackall et al., 2006; Seppenwoolde et al., 2002; Suh et al., 2008] and from one breathing cycle to the next (known as inter-cycle variability) [Blackall et al., 2006; Von Siebenthal et al., 2007c].

2.3.1 Inter-cycle Variability

Variations between one breathing cycle and other breathing cycles are referred to as inter-cycle variabilities. This can either be differences in the amplitude or trajectory from one breathing cycle to the next [Keall et al., 2006; Blackall et al., 2006], or more systematic respiratory drift which can occur during longer imaging or treatment sessions [Von Siebenthal et al., 2007a].

Simple variations in magnitude and shape of respiratory motion trajectories occur between consecutive breathing cycles and to a much larger degree for breathing cycles separated by a long time. Kini et al. [2003] investigated these variations by looking at respiratory traces obtained using the Varian Real Time Position Management (RPM) system, which derives a 1D signal by tracking infra-red surface markers placed on the patient's chest and abdomen. An example of such a trace on two separate treatment days is reproduced in Fig. 2.2. Slight variations in the magnitude and shape of the respiratory trajectory can be observed from breathing cycle to breathing cycle. Over longer periods, such as here on different treatment days, significant changes to the motion pattern can occur.

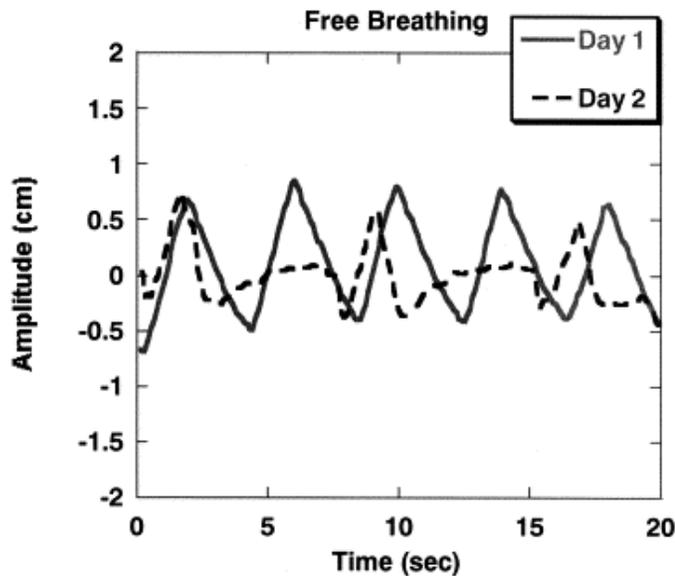


Figure 2.2: Respiratory traces obtained using the Varian RPM in a patient on two different treatment days. Figure reproduced from [Kini et al., 2003].

In addition to changes in the respiratory pattern such as observed by Kini et al. [2003], respiratory motion may also undergo a continuous drift. Von Siebenthal et al. [2007a] studied the respiratory motion of the liver over extended periods of time using 4D MR. The authors found that the liver can undergo significant drift in the time scale of a radiotherapy

treatment due to bowel movements or general muscle relaxation. It was found that the average drift in the exhale position of 11 healthy volunteers amounted to 3.1 mm in 20 minutes, however the maximum drift ranged up to 12.8 mm.

It is important to note that one-dimensional motion surrogate signals such as, for example, the one shown in Fig. 2.2 or signals obtained using a respiratory bellows [McClelland et al., 2013] or an MR pencil beam navigator [Köhler et al., 2011] can be misleading. A surrogate signal may have identical values at different time points while the internal organs have significantly different configurations. Blackall et al. [2006] investigated this effect using dynamic MR imaging and motion modelling (see Section 3.3) in patients and found that lung surface landmarks at equal surrogate values may vary by up to 3.9 mm for different breathing cycles (but during the same imaging session).

2.3.2 Intra-cycle Variability

Hysteresis occurs when a location (such as a tumour) follows a different path during inhalation than during exhalation. Seppenwoolde et al. [2002] found that hysteresis in the lungs, when present, was largest in the sagittal plane, but could also be observed in other planes. They observed hysteresis in the trajectories of half the patients studied, amounting to a 1 mm to 5 mm separation of the trajectories during inhalation and exhalation. To illustrate the large variability in trajectories that tumours can take the authors plotted the trajectories of 21 tumours in the coronal and sagittal planes. This diagram is reproduced in Fig. 2.3. Significant hysteresis can also occur in the liver [Von Siebenthal et al., 2007a] and the heart [Nehrke et al., 2001].

As for the inter-cycle variability, caution must be exercised when measuring respiratory cycles using a 1D surrogate signal, as similar values may correspond to different organ configurations. Blackall et al. [2006] also investigated this effect for intra-cycle variabilities and found that lung surface landmarks may vary by up to 5.2 mm in patients between inhaling and exhaling trajectories of the same breathing cycle. The effect is most pronounced close to the inhale state.

2.4 Breath-holding

Breathing can be consciously interrupted for short periods of time for a breath-hold. Breath-holding is one of the simplest solutions to respiratory motion as will be discussed in more detail in Section 3.1.1. Normal, healthy subjects can usually comfortably hold their breath for maximally 30 seconds [Scott et al., 2009]. However, breath-holding can often be problematic for patients and breath-hold durations may be limited to shorter periods [Scott et al., 2009; Zech et al., 2004]. Breath-holds can be performed at any respiratory posi-

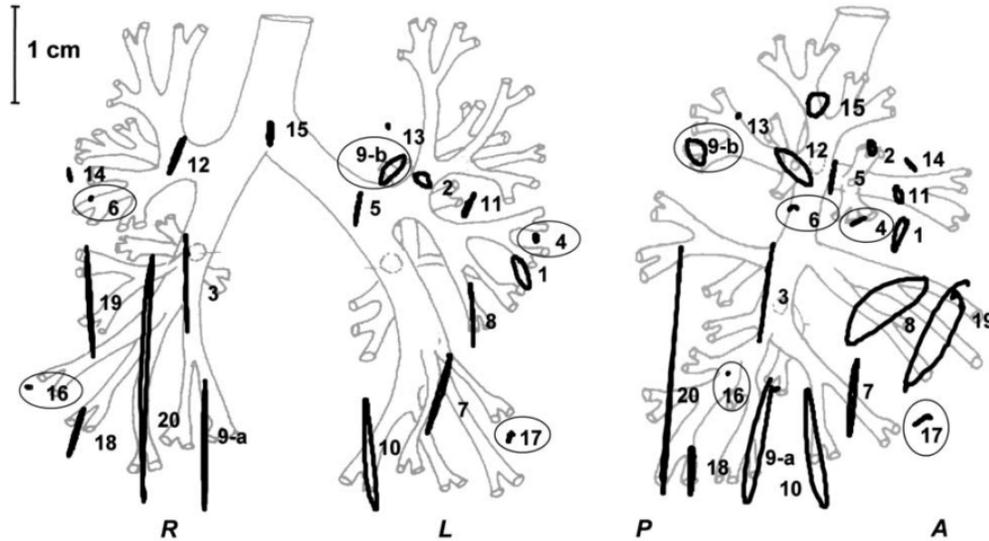


Figure 2.3: Orthogonal projections of the trajectories of 21 tumours on (left) the coronal and (right) the sagittal plane. The tumours are displayed at their approximate location and tumours that were attached to bony structures are circled [Seppenwoolde et al., 2002].

tion, but most commonly either end-exhale, or end-inhale are used because patients can be coached to achieve those states with a reasonable degree of reproducibility [Keall et al., 2006; Blackall et al., 2006]. Inhale states may be beneficial for some applications, such as radiotherapy, because in this configuration healthy lung tissue is less likely to be in the primary radiation beam. However, it has been shown that self-determined inhale breath-holds are less reproducible [Blackall et al., 2006; Liu et al., 1993]. However, even exhale breath-holds can be characterised by significant variations. Liu et al. [1993] attempted to quantify variability in subsequent, self-determined exhale breath-holds, and found that the diaphragm position varied 8.3 mm on average between breath-holds. Lastly, it is known that the predominant respiratory muscles can be different for breath-holding compared to free breathing [Keall et al., 2006; Blackall et al., 2006]. This may have implications for applications where data acquired during breath-holding is used to predict free breathing motion states.

2.5 Problems Caused By Respiratory Motion

Respiratory motion is a complicating factor in a range of applications. They can be broadly divided into medical imaging techniques where the respiratory motion can cause various kinds of artefacts, and medical interventions where respiratory motion can cause the intended target volume to deviate from pre-treatment planning. In the following, a selection

of applications from each category will be discussed.

2.5.1 Interventions

radiotherapy (RT) is often used as part of curative or palliative treatments of various cancers. In RT ionizing radiation, which causes cell death, is used to irradiate tumours in order to eliminate them or control their growth. In order to maximise the dose the tumour receives and at the same time minimise the irradiation of surrounding healthy tissue, typically RT treatments are planned based on 3D volumes obtained using anatomical imaging techniques such as computed (X-ray) tomography (CT) [Lichter and Lawrence, 1995] or functional imaging techniques such as PET. When treating lesions in the thorax or abdominal organs, respiratory motion can cause the tumour to move out of the planned target volume, which in the worst case causes healthy tissue to be irradiated instead of the tumour. To account for this, in the RT planning stage, safety margins are typically placed around the target volume [Seppenwoolde et al., 2002]. More accurate target volumes with smaller safety margins allow for higher doses, which in turn have been shown to increase the survival rate for various cancers [Keall et al., 2006]. Thus, accurate knowledge of the tumour respiratory motion is highly beneficial and many of the studies outlined above aim to characterise this motion.

In addition to the advent of non-invasive radiation therapy in the last few decades there has been a shift from open surgeries to minimally invasive surgeries, as they lead to much reduced damage of surrounding tissues which results in lower patient mortality, faster recovery and shorter hospital stays [Peressutti, 2014]. Conventionally, a 3D preoperative plan, based on images acquired before the procedure, is registered to the physical space of the patient in the operating room. In interventions on the heart, such as atrial ablation or balloon dilation of pulmonary valves [Tzifa et al., 2010; Klemm et al., 2007], respiratory motion causes significant challenges. Similar to RT, it may cause a misalignment between the anatomy and the guidance information acquired pretreatment [Hawkes et al., 2005; King et al., 2009a].

Another minimally invasive technique which is currently being studied for the treatment of certain abdominal tumours is high-intensity focused ultrasound (HIFU) ablation. In HIFU acoustic energy is focused on the cancerous tissue in order to raise its temperature and cause ablation. However, applying HIFU to the organs of the upper abdomen is challenging as respiratory motion can cause significant displacements of the organs [Ries et al., 2010]. Gating approaches have been proposed but may lead to very lengthy treatment times [Rijkhorst et al., 2011]. To allow for more accurate tracking of tumours it has been proposed to combine HIFU with MR guidance [Hynynen et al., 1996]. As well as visualising the anatomy and tumour in real-time, such a system can produce real-time temperature

maps based on MR phase imaging [Ries et al., 2010], which allow very accurate dosage of the deposited energy. However, in addition to the aforementioned target motion, respiratory motion in this application results in undesired phase variations which can lead to significant temperature artefacts [Köhler et al., 2011; Ries et al., 2010].

2.5.2 Imaging of Anatomy with Respiratory Motion

In addition to the problems mentioned above for interventions, respiratory motion of the thorax and upper abdomen also affects the acquisition of images using various modalities. This may limit their utility for planning treatments such as the ones outlined above, but also their clinical utility in general.

MR imaging is in wide clinical use for medical diagnosis, staging of disease and follow-up. The modality can provide high soft tissue contrast without exposure to ionizing radiation [McRobbie et al., 2006]. However, in MR imaging of organs in the thorax or upper abdomen respiratory motion can lead to inconsistencies in k-space which in turn can cause artefacts such as intensity loss, phase inconsistencies, blurring or ghosting, all of which can reduce the images' clinical utility [Ehman and Felmlee, 1989; McRobbie et al., 2006]. This can be a limiting factor for numerous applications, such as for example diagnosing lesions in the liver [Zech et al., 2004] or imaging of the coronary arteries [Nehrke et al., 2001; Wang et al., 1995]. One widely used solution to the problem of respiratory motion is acquiring the scan during a breath-hold. However, for many applications scans of sufficient signal-to-noise ratio or resolution cannot be completed within the comfortable breath-hold period of 20-30 seconds for a healthy volunteer [Scott et al., 2009; Zech et al., 2004]. Acquiring a respiratory signal using, for example, a respiratory bellows or an MR pencil beam navigator allows the use of gating techniques whereby imaging data are only acquired in a predefined window of signal values [Ehman et al., 1984; McRobbie et al., 2006]. However, such approaches often significantly prolong acquisition times [Ehman and Felmlee, 1989] and may lead to images of inferior quality compared to breath-hold images [Katayama et al., 2001].

Respiratory motion poses particular challenges in PET. PET allows the visualisation of functional processes in the body by detecting photon pairs, which are generated when positrons emitted by a specific radioactive tracer annihilate with electrons in the body. The tracer is bound to a biologically active molecule and is injected into the bloodstream pre-treatment. At present the most commonly used clinical tracer is fluorodeoxyglucose (FDG), which is fluorine-18 attached to an analogue of glucose. The technique is of particular interest for visualising tumours by measuring their metabolic activity. In PET imaging of the thorax or abdomen respiratory motion poses a significant problem as typical clinical PET acquisitions currently take between two to ten minutes per bed position [Thielemans et al., 2011; Grimm et al., 2013]. Therefore, breath-holding techniques are not feasible and

gating techniques lack wide clinical acceptance due to low SNR or excessive scan times [Grimm et al., 2013]. Respiratory motion may lead to blurring of lesions in PET, which may lead to reduced detectability of small tumours, errors in creating planning volumes for radiotherapy treatments and incorrect standardised uptake value (SUV) calculations [Visvikis et al., 2006; Nehmeh et al., 2008]. Incorrect SUV measurements can in turn adversely affect diagnosis and tumour staging [Thielemans et al., 2011; Liu et al., 2009].

Attenuation correction is a method for correcting the PET data for the effects of the different types of tissues the photons pass through on their way to the detectors. Dense tissue, like bones, will absorb a larger part of the photons whereas soft tissue, like lungs, will absorb much less. Images reconstructed without attenuation correction will, thus, show apparently greater activity in areas of soft tissue as compared with dense tissue. Typically, attenuation correction is performed based on a single CT scan. However, this scan represents a snapshot of a single respiratory state, and hence artefacts will occur if PET data is acquired at a different respiratory state.

2.6 Discussion

Respiratory motion poses significant challenges in a wide range of medical treatments and imaging techniques. Furthermore, because treatments increasingly rely on robust imaging those problems have become strongly interconnected.

Accurate imaging and correction for this motion is crucial as it may directly influence the survival of patients, for example following minimally invasive treatments, or due to misdiagnosis of cancer because of imaging artefacts. Due to the high variabilities respiratory motion may exhibit, both between inspiration and expiration (intra-cycle variabilities) and between separate breathing cycles (inter-cycle variabilities), this is a very challenging problem.

In the following chapter existing solutions to the problems that respiratory motion poses, and methodologies to better understand respiratory motion will be discussed.

Chapter 3

Literature Review

Contents

3.1	Correcting for Respiratory Motion	23
3.1.1	Breath-Holding	23
3.1.2	External Gating Techniques	23
3.1.3	Self-Gating Methods	26
3.1.4	Motion Correction Using Another Modality	27
3.2	Imaging Respiratory Motion	30
3.2.1	Dynamic Imaging	30
3.2.2	Slice-Stacking	31
3.3	Patient-Specific Motion Modelling	31
3.3.1	Model Calibration	32
3.3.2	Model Formation	33
3.3.3	Model Application	37
3.4	Manifold Learning	39
3.4.1	Introduction	39
3.4.2	Manifold Learning For Understanding Physiological Motion	40
3.5	Manifold Alignment	41
3.5.1	Methods for Alignment	43
3.5.2	Types of Supervision	44
3.6	Discussion	45

In this chapter some important works from the fields of respiratory motion correction, manifold learning and manifold alignment will be discussed. First, in Section 3.1 I will talk about strategies for motion correction in various imaging modalities and treatments. In Section 3.2 possibilities to directly or indirectly image respiratory motion using fast dynamic imaging modalities will be discussed. Another approach to respiratory motion,

motion modelling, will be examined in Section 3.3. Finally, the last two sections of this Chapter (Sections 3.4 and 3.5) will examine manifold learning and manifold alignment techniques with a particular focus on applications in medical imaging.

3.1 Correcting for Respiratory Motion

For a variety of medical applications such as the diagnosis of illnesses or the planning of treatments, physicians are interested in imaging the thorax or abdomen. Often this is not possible because acquisition times are long when compared to typical breathing cycle lengths. In the following a number of strategies for correcting for respiratory motion are examined.

3.1.1 Breath-Holding

As mentioned in Chapter 2, for simple image acquisitions breath-holding can be a feasible solution. However, breath-holding can often be problematic for patients [White et al., 2009], and acquisitions are limited to maximally 30 seconds [Scott et al., 2009; Zech et al., 2004]. While some simple imaging techniques can be performed during a single breath-hold, more sophisticated techniques require the combination of multiple breath-hold scans. For example, Nehmeh et al. [2007] investigated combining PET data acquired during nine 20 second breath-hold scans at a deep inhale position.

Combining multiple breath-holds has also been proposed for various treatments. Mah et al. [2000] investigated deep inspiration breath-holding for the treatment of seven patients with thoracic tumours using radiotherapy. Deep inspiration was chosen because it expanded the patient's lungs to maximum volume thus driving healthy lung tissue out of the primary radiation beam. A similar methodology has also been applied to MR-guided HIFU treatments of the liver. Okada et al. [2006] performed sonifications during 46 subsequent 20 second breath-holds in a single patient with hepatocellular carcinoma. The respiratory position was monitored using a respiratory bellows and the temperature elevation was visualised using phase MR.

For the reasons discussed in Section 2.4, combining multiple breath-hold gates may not be optimal under some circumstances as they may not be reproducible. Furthermore, breath-holding may only be feasible for a subset of patients.

3.1.2 External Gating Techniques

Respiratory gating techniques measure the respiratory motion using some simple surrogate signal, which can be easily acquired during the treatment or imaging session and has a strong correlation with the internal motion. This signal is then used to control treatments;

for example it can be used to decide when to turn on the radiation beam in radiotherapy. In imaging it can be used to restrict the data collection to a single gating window in the respiratory cycle. Advanced techniques combine multiple gating windows to achieve either better quality or time-resolved images. In the following, first a few of the most common methods to obtain a respiratory surrogate signal are discussed before reviewing a number of different gating methods. In particular, in Section 3.1.4, a number of techniques will be presented which have used gating to correct for motion in simultaneous PET/CT and PET/MR.

Commonly Used Devices for Measuring Respiratory Motion

In this section the different ways of making physical measurements to derive a respiratory trace for gating and later for motion modelling are reviewed.

In MR image acquisition the most common surrogate data proposed has been the MR navigator echo, also called a pencil beam navigator. This involves a small column of magnetisation being excited to measure the position of a region of tissue over time [Danas et al., 1997]. They have been most commonly used to track the head-foot translation of the right hemi-diaphragm.

The respiratory bellows is an alternative means of measuring respiratory position [Santelli et al., 2011]. This consists of an air filled bag, which is wedged between the subject's abdomen or chest and a firm surface such as an elasticated belt. The motion of the abdomen or chest during respiration causes air to be expelled from the bellows and a sensor measures the flow of the air.

A spirometer measures the air flow to and from the lungs, and is commonly used for testing pulmonary function. One problem with using spirometry as a surrogate signal is that there can be considerable drift in the spirometry signal due to instrumentation errors and/or escaping air [McClelland et al., 2013].

A common means of acquiring respiratory surrogate data has been to track the motion of one or more points on the surface of the chest or abdomen. There are various commercially available products which implement such solutions, for example the Varian real time position management (RPM) system¹.

Techniques Using a Single Gate

Gating methods for radiotherapy were first demonstrated in Japan by Ohara et al. [1989]. The authors controlled the beam activation using a respiratory signal measured by a spirometer for lung cancer patients. The beam was only triggered when the patient was in a

¹<https://www.varian.com/oncology/products/motion-management-verification/rpm-respiratory-gating>

predefined exhale gate. Gating has also been proposed for MR-guided radiotherapy in an experimental hybrid MR scanner/linear accelerator [Stam et al., 2012]. However, the method has not yet been demonstrated on patients.

A similar method was proposed recently for the gating of MR-guided HIFU [Auboiroux et al., 2014]. The gating was performed based on signals extracted automatically from images acquired using a digital MR-compatible camera.

Respiratory gating for MR imaging was first proposed by Ehman et al. [1984]. The authors investigated three different methods for obtaining respiratory traces: Two devices measuring the temperature of inhaled and exhaled air, and a respiratory bellows. The authors found that the latter led to the most accurate results, and respiratory bellows are still in wide clinical use today.

The first PET gating technique was proposed by Nehmeh et al. [2002]. The authors measured the respiratory trace using the Varian RPM system and divided the respiratory cycle into a varying numbers of bins. The PET data was sorted into the appropriate bins but only the data from the first bin was reconstructed. That is, the gates were not combined. This has the advantage of significantly reducing the blurring resulting from respiratory motion [Nehmeh et al., 2002], but results in reduced PET counts and thus leads to noisier images [Nehmeh et al., 2002; Bai and Brady, 2009].

Combining Multiple Gates

In particular for imaging techniques conventional gating techniques such as described above lead to very inefficient protocols because for a large fraction of the time no data is acquired or the acquired data is discarded. Therefore, a large number of techniques have been proposed to combine data from multiple gates into a single image.

Examples include a recently proposed technique by Aitken et al. [2014] for 100% efficient MR-based coronary angiography. The authors bin data based on translation estimates obtained from a low-resolution 2D MR navigator. The data from the bins are then combined by affinely registering them to a reference bin.

A similar methodology has also been proposed for PET imaging. Dawood et al. [2006] performed a gating of list-mode data into eight bins using a respiratory bellows signal and reconstructed each bin separately without attenuation correction. In a next step an optical flow registration algorithm was applied to transform seven of the gates to the respiratory position of the exhale gate. Finally, the data were summed to produce the final motion corrected PET images. Bai and Brady [2009] proposed a very similar method which also included respiratory motion in the attenuation correction.

Time-Resolved Imaging

Lastly, in some cases the desired output is not a static image but a series of images informing the clinician about organ motion during the respiratory cycle. Such information can be crucial for planning of interventions such as radiotherapy. In that case, the data collected in different gates is not combined into a single gate but is either used directly, or in some instances, used to reconstruct multiple improved gates.

It has been shown that when including dynamic information obtained using 4D CT into treatment planning better target coverage can be achieved [Colgan et al., 2008]. 4D CT was first proposed by Vedam et al. [2002]. The authors acquired an external respiratory signal using the Varian RPM system. In addition, a respiratory phase signal capable of separating inhalation from exhalation was derived from the original infra-red signal. The images were acquired in helical mode with a very slow couch speed, such that adjacent images were only 0.5 mm apart. They were then sorted into 8 predefined bins (4 inhalation, 4 exhalation) and each bin was combined into a volume. Low et al. [2003] proposed a simple gating method for 4D CT images of the lungs. Scans were acquired in cine mode, in which the couch position was kept constant for a series of 15 scans before the patient was shifted to the next couch position. This allowed for higher sampling of the respiratory states. Simultaneously, tidal volume was measured using a digital spirometer. The images were then retrospectively sorted by matching tidal volumes across the different couch positions. A 4D approach has also been proposed for four-dimensional cone-beam CT (4D CBCT) imaging [Dietrich et al., 2006].

An important drawback of most time-resolved imaging techniques based on gating is that they typically only acquire partial data at one or more predefined gating windows and need to wait until a certain respiratory state is revisited multiple times before a satisfactory image can be produced. This leads to an averaging of the images over a number of respiratory cycles, which in turn may lead to various artefacts [Yamamoto et al., 2008].

3.1.3 Self-Gating Methods

Another approach to gating which is relevant to this thesis is to extract a respiratory trace for gating directly from the acquired data itself. Such approaches are commonly referred to as self-gating, or data driven techniques. Most commonly in these techniques the gating is performed retrospectively, i.e. after the data acquisition.

For PET data Visvikis et al. [2006] measured the time-activity curves in regions of interest that included a border between an organ with high uptake and one with low uptake. This then allowed the amplitude and phase of each frame of the acquisition to be found. A similar approach was proposed by Bundschuh et al. [2007], where the PET data were first binned into short time bins. In order to derive a trace, in the next step the centre of

mass in the S-I direction in a manually defined volume of interest was calculated for the reconstruction of each bin. Lastly, a self-gating approach for PET data using linear dimensionality reduction, i.e. principal component analysis (PCA), was proposed by Thielemans et al. [2011]. This will be discussed in more detail in Section 3.4.

In the MR literature a frequently used method for self-gating is the use of acquisition sequences which oversample the centre of k-space. The changes in the DC component of the acquired images can be used to extract a respiratory trajectory [Pipe et al., 1999]. Recently, Cruz et al. [2015] proposed a self-gating technique for abdominal MR imaging. The data were acquired using a 3D radial phase encoding sequence [Prieto et al., 2010]. The respiratory trace derived from the k-space centre was used to bin the data into gates, which were subsequently corrected for motion using non-rigid registration.

Self-gating techniques have also been employed for time-resolved imaging. Sundar et al. [2009] proposed a general respiratory gating technique and applied it to, among other applications, the gating of 3D ultrasound images of the liver. The respiratory phase was estimated by spectral analysis of the images.

Two self-gating approaches based on non-linear dimensionality reduction have also been proposed for the gating and compounding of 3D ultrasound images [Wachinger et al., 2011] and for 4D CT reconstruction [Georg et al., 2008]. These approaches will be further discussed in Section 3.5. In Chapter 6, it will be demonstrated how manifold alignment can be used for the self-gating and compounding of multiple 3D US views.

3.1.4 Motion Correction Using Another Modality

In PET, motion fields can be extracted directly from the reconstructed gates themselves as described above [Bai and Brady, 2009; Dawood et al., 2006]. However, this approach may lack robustness for specific radiotracers with low background uptake [Würslin et al., 2013]. Therefore, PET data is often acquired together with another modality, such as 4D CT or 4D MR, and motion fields are extracted from those images using registration techniques. Such motion fields may then either be used to retrospectively transform a number of reconstructed PET gates to a common reference in a process called reconstruct transform-average (RTA) as has been done above in Dawood et al. [2006] and Bai and Brady [2009], or they may be incorporated directly into the PET reconstruction process for a motion corrected image-reconstruction (MCIR) [Qiao et al., 2006].

The majority of the literature on PET motion correction using 4D CT is based on the MCIR motion correction approach. For example, Qiao et al. [2006] and Lamare et al. [2007] both proposed to extract non-rigid 3D motion fields from 4D CT images using B-spline registration and incorporate these directly into the PET reconstruction process.

CT imaging comes at the cost of increased radiation exposure to the patient [Beyer et al.,

2000]. MR imaging is a promising alternative to CT because of its non-ionising nature, and its good soft tissue contrast. The recent emergence of hybrid PET/MR systems has opened up the possibility of deriving the motion fields needed for PET motion correction from simultaneously acquired MR images [Tsoumpas et al., 2010; Würslin et al., 2013].

Employing such a hybrid system, Grimm et al. [2013] used a golden radial phase encoding 3D sequence to acquire MR data simultaneously with the PET data. Similar to the works described in Section 3.1.3, a respiratory trace was extracted from the k-space centre. This signal was subsequently used to bin the MR k-space data, as well as the PET list-mode data. In a next step, the MR as well as the PET data from each bin were reconstructed. Lastly, motion fields were derived from the MR data using non-rigid registration and used to warp the reconstructed PET gates to the exhale motion state using a RTA approach.

The above approach may however be suboptimal when correcting for respiratory motion in the lungs. Current 3D MR imaging technology is limited by the lack of contrast in dynamic MR images of the lungs. The rapidly decaying MR signal from the lung parenchyma makes the lung appear dark in dynamic 3D images [Robson et al., 2003], prohibiting robust motion field estimation. Since lung motion is not homogeneous, accurate motion fields cannot be simply interpolated between the lung borders [Ding et al., 2009]. Furthermore, dynamic 3D MR images suffer from low image resolution and relatively long acquisition times, which can lead to motion blurring, further limiting the accuracy of motion estimation. Dynamic 2D MR scans, on the other hand, can be acquired in a shorter time frame, have excellent in-plane resolution and, when acquired at systole, have high contrast in the lungs due to the in-flow of previously unexcited blood into the slice. However, they lack the coverage of 3D scans. Examples of dynamic 2D and 3D MR acquisitions obtained using T1-weighted gradient echo sequences on a 3T MR scanner are shown in Figure 3.1.

Combining the favourable contrast and resolution properties of 2D slices with the coverage of 3D volumes is possible by acquiring 2D slices multiple times in a slice-by-slice fashion and then combining them using a slice stacking technique [Dikaïos et al., 2012; Von Siebenthal et al., 2007a].

Dikaïos et al. [2012] proposed continuously acquiring coronal 2D slice-by-slice data from varying slice positions and then reconstructing 3D volumes for PET motion correction using a simple slice stacking scheme. For each slice the spatially neighbouring slice that maximised the normalised mutual information between the slices was chosen. This process was repeated for the newly found slice until the whole volume was reconstructed. The exact gating strategy was not discussed, but was presumably performed using an external signal such as a respiratory bellows. Both MCIR and RTA motion correction were evaluated on synthetic PET/MR data, with the MCIR giving slightly better results.

Würslin et al. [2013] used a spoiled gradient echo imaging sequence to obtain sagittal

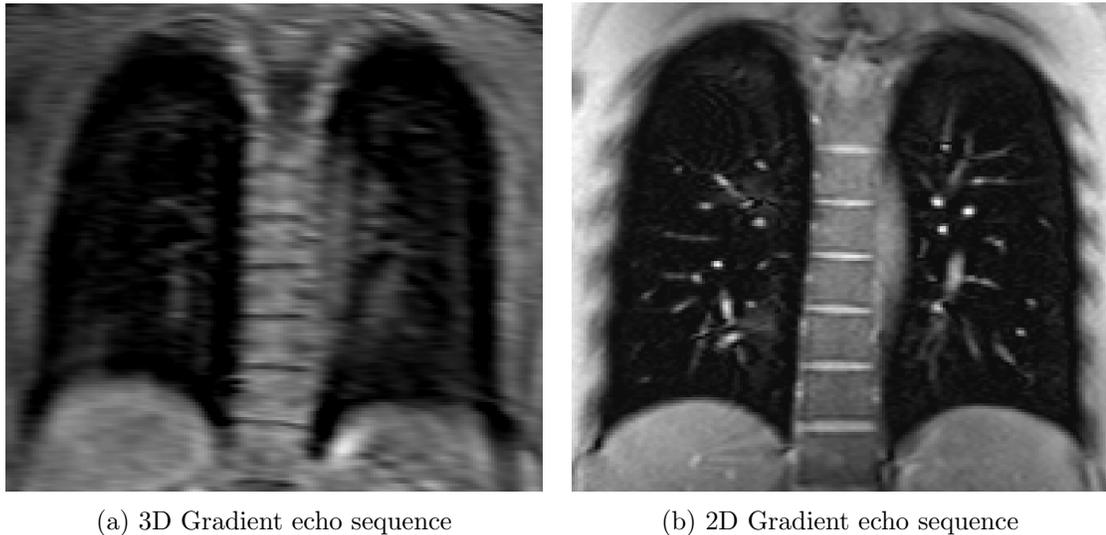


Figure 3.1: Comparison of two MR acquisitions covering the lungs: (a) a coronal slice of a 3D T1-weighted gradient echo sequence; b) Coronal slice of a 2D T1-weighted gradient echo sequence. Here, the large pulmonary arteries are visible due to the in-flow of previously unpolarised blood.

2D MR slices. Each slice position was acquired 12 times before moving on to the next one. In addition to the slice data, the authors acquired a 1D MR navigator echo [Ehman and Felmlee, 1989] prior to each image acquisition. They then used this to bin the 2D MR slices as well as the PET data. The MR data was reconstructed by stacking all slices in the same bin. That is, the MR image closest to the bin's mean respiratory position was inserted into the corresponding 3D volume, and the remaining 11 images were discarded. Using a demons registration algorithm non-rigid motion fields were derived from these gated 3D volumes, which were then used to correct the reconstructed PET gates for motion in a RTA fashion.

In Chapter 5, a manifold alignment-based technique for PET/MR motion correction will be proposed and compared against the two techniques outlined above.

Although MR has great potential for respiratory motion correction of PET list-mode data it is important to note that attenuation correction is much harder to perform using MR imaging data. In contrast to CT images, the image intensities in MR are not related to the electron density. A number of approaches for MR based attenuation correction have been proposed and are discussed in Wagenknecht et al. [2013].

3.2 Imaging Respiratory Motion

For many applications high-quality, time resolved images or volumes of organs in the thorax are crucial. Some modalities can be acquired with sufficient temporal resolution that motion artefacts become negligible. However, this usually comes at the cost of diminished spatial resolution, coverage or image quality. If dynamic imaging is not possible, gated imaging (such as, for example, 4D CT) often provides a feasible alternative.

3.2.1 Dynamic Imaging

In order to gain understanding of how organs move due to respiration, for accurate treatment planning or for the retrospective motion correction of simultaneously acquired modalities, it is often desirable to obtain dynamic images depicting the respiratory motion. US is a comparatively cheap imaging modality which can be used to obtain dynamic images in 2D or 3D. However, US suffers from a number of artefacts such as speckle noise and shadowing, and has limited utility for applications where high image detail is required. Furthermore, imaging of the lungs is not possible with US, because sound waves are significantly attenuated by air. Another commonly used technique is X-ray fluoroscopy. While offering high temporal resolution, this modality is limited to 2-dimensional projections, and is based on ionizing radiation. In addition, image quality may be poor, in particular when low X-ray doses are used.

MR is highly versatile and various acquisition speeds, fields of view, and contrasts can be achieved by employing and combining different pulse sequences. The echo planar imaging (EPI) sequence allows extremely fast 2D slice acquisitions with around 50-100 ms per slice [Stehling et al., 1991]. In a more recent work the EPI sequence was modified for imaging the left ventricle. Using a small field of view covering only the ventricle, slice acquisition times down to 20 ms were achieved [Nayak et al., 2001]. However, these sequences typically suffer from low signal-to-noise ratio, low resolution, and image artefacts [Bushberg et al., 2001]. It is also possible to acquire high-resolution slices of the thorax using gradient echo (GRE) sequences in less than 200 ms per slice. For example the slice shown in Fig. 3.1b was acquired using a T1-weighted GRE sequence in approximately 180 ms. Furthermore, it is also feasible to acquire 3D volumes with limited blurring due to respiratory motion. A coronal slice of a volume also acquired using a T1-weighted gradient echo sequence in approximately 600 ms is shown in Fig. 3.1a. However, clearly the image quality is much worse compared to the 2D slice.

3.2.2 Slice-Stacking

Attempts have been made to combine the high-resolution and contrast properties of 2D MR slices with the coverage of 3D volumes, by stacking such slices acquired at different positions. A few such methods, which rely on gating, have already been discussed in Section 3.1.4. Slice-stacking can also be used to reconstruct image sequences, which do not average the motion over several motion cycles, but attempt to give a true representation of organ movement due to respiratory motion including all variations over a period of time.

A closely related work is that by Von Siebenthal et al. [2007a], who proposed a 4D MR imaging technique based on slice stacking for quantifying errors in radiotherapy treatments. The data was acquired in an interleaved 2D MR slice-by-slice sequence, where a navigator slice at a constant slice position was acquired before each of the data slices, which imaged a different plane each time. Manually selected image features derived from the navigator slices were then matched retrospectively to find data from all different planes which were acquired at similar respiratory positions. Using this information, a 3D volume was reconstructed around each of the acquired 2D MR slices in the original acquisition order, resulting in long 4D MR sequences. The drawback of this method is that such image acquisitions are very time-consuming as 50% of the data are not used for reconstruction (i.e. the navigator slices). In the paper, in order to reconstruct a 30 minute sequence, 60 minutes of imaging had to be performed.

In Chapter 5, a 100% efficient slice stacking technique based on manifold alignment will be proposed.

3.3 Patient-Specific Motion Modelling

A different approach to address the problem of respiratory motion is the use of patient-specific motion models, which is the topic of Chapter 6 of this thesis. As opposed to gating approaches the data is not directly sorted or binned based on a respiratory surrogate signal. Rather, motion measurements are derived from imaging data and a direct or indirect relationship between the motion fields and a simultaneously acquired surrogate signal is formed. After the model is built it is possible to obtain motion fields by measuring only the surrogate [McClelland et al., 2013]. The motion models discussed in this Section are summarised in Table 3.1. Note that patient-specific motion models are distinct from population-based motion modelling approaches, where motion estimates are derived from a previously formed motion atlas [Ehrhardt et al., 2011; Arnold et al., 2011]. These will not be further discussed in this thesis.

Subject-specific motion models have been proposed to investigate the respiratory motion of various organs and for various applications. The first uses of motion models were mostly

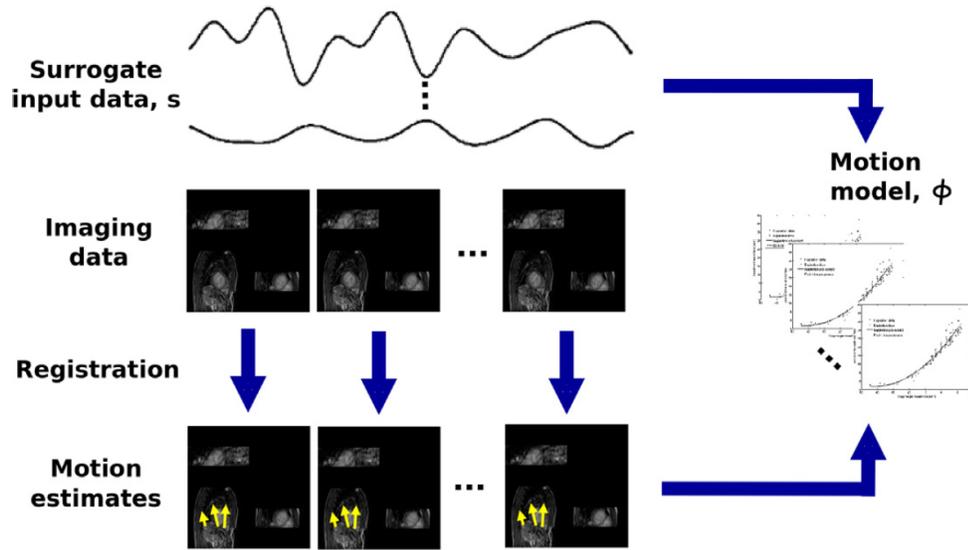


Figure 3.2: Schematic illustration of motion model formation and application. Surrogate data is acquired at the same time as some imaging data representing the motion of interest. The motion is estimated from the imaging data, typically using registration, and the motion model establishes a correspondence between the surrogate data and the motion. In the model application stage, motion estimates can be generated from previously unseen surrogate data (not shown). The figure is reproduced from McClelland et al. [2013].

focused on the reduction of motion artefacts in the acquisition of MR images [Wang and Mahadevan, 2009; Manke et al., 2002, 2003; White et al., 2009], and target tracking in radiotherapy treatments [Schweikard et al., 2000, 2005; Gao et al., 2008; Isaksson et al., 2005]. In recent years, motion modelling has been proposed for a wide range of different imaging techniques and image-guided treatments, in particular the acquisition of PET/CT [Fayad et al., 2010] and PET/MR [King et al., 2011, 2012], CT acquisitions [McClelland et al., 2006; Zhang et al., 2010], image-guided cardiac interventions [King et al., 2009a,b; Peressutti et al., 2012, 2013], and image-guided HIFU interventions [Rijkhorst et al., 2011].

Motion models typically consist of three distinct stages: (1) model calibration, (2) model formation, (3) model application. The formation and application of a typical model are summarised in Fig. 3.2. In the following each of the steps will be discussed in more detail.

3.3.1 Model Calibration

In a first step, imaging data are acquired prior to the treatment or imaging session from the anatomical region of interest and motion is estimated from them. The motion estimation is commonly performed using affine [Manke et al., 2003; King et al., 2009b,a; Peressutti et al., 2013] or non-rigid registration [Ablitt et al., 2004; Zhang et al., 2010; McClelland et al., 2006; Gao et al., 2008; King et al., 2011; Rijkhorst et al., 2011; King et al., 2012].

Depending on the way the motion was derived, the motion can internally be described by the affine transformation parameters [Manke et al., 2003; King et al., 2009b,a; Peressutti et al., 2013] or by full deformation fields [Zhang et al., 2010; White et al., 2009; Fayad et al., 2010]. Frequently, if the motion fields were derived using a B-spline based registration, the motion is directly represented by the displacements of the B-spline control point grid [Ablitt et al., 2004; McClelland et al., 2006; Gao et al., 2008; King et al., 2011; Rijkhorst et al., 2011; King et al., 2012].

In some instances, in particular for radiotherapy treatments, only the motion of a target, e.g. the tumour, is modelled. In that case it is sufficient to model just the displacements of (a number of) points of interest, which can be automatically tracked in the input images [Low et al., 2005; Schweikard et al., 2000; Isaksson et al., 2005].

Simultaneously with the imaging data, simpler surrogate data are acquired which are correlated with the motion and can also easily be obtained or approximated during the treatment. In the simplest case the surrogate data are one-dimensional signals such as already discussed for gating (see Section 3.1.2). Examples include MR pencil beam navigators [Manke et al., 2002; King et al., 2011], traces derived from IR marker tracking [McClelland et al., 2006; Schweikard et al., 2000] and spirometers [Low et al., 2005]. In other instances a simpler signal is derived from high-dimensional imaging data, most commonly by tracking the motion of the diaphragm in the S-I direction in the images [Blackall et al., 2006; King et al., 2009a,b; Rijkhorst et al., 2011].

As will be discussed in more detail in the next section motion models vary in their ability to capture respiratory inter- and intra-cycle variabilities, which were discussed previously in Section 2.3. In an attempt to model more of these variabilities, in recent years there has been a trend towards the use of more complex surrogate data. Some works simply increase the number of measured surrogate signals. For example, Manke et al. [2003] used three MR pencil beam navigators at different anatomical locations in order to increase the amount of captured variability, and Zhang et al. [2010] extracted two temporally shifted respiratory traces from 4D CT data. Other works have used more complex respiration surrogates such as chest surface data [Ablitt et al., 2004; Fayad et al., 2010; Grau et al., 2007], real-time imaging data such as US images [Peressutti et al., 2013] or fast 2D MR images [King et al., 2012], or partial k-space data [White et al., 2009].

3.3.2 Model Formation

After the motion has been estimated from the imaging data and suitable surrogate data has been acquired, in a next step a model is formed by finding a correspondence between the two. The correspondence between surrogate data and the motion can be estimated in two ways. First, the motion representation can be expressed *directly* as a function of the

surrogate.

An early paper using a simple form of motion modelling is Wang et al. [1995]. The authors were the first to propose a simple linear model, where the displacement of the diaphragm (i.e. the surrogate data) was related to the displacements of cardiac landmarks through a fixed linear relation [McClelland et al., 2013].

A common approach to establish a direct correspondence model is to express each motion descriptor (e.g. affine transformation parameter or B-spline control point) as a function of the surrogate(s). Some possible choices from the literature are linear functions [Manke et al., 2002; Ablitt et al., 2004; Low et al., 2005; Zhang et al., 2010; Manke et al., 2003; Gao et al., 2008; Fayad et al., 2010; Schweikard et al., 2000], higher degree polynomials [Blackall et al., 2006; King et al., 2009a,b, 2011; Rijkhorst et al., 2011] and cyclic B-spline functions [McClelland et al., 2006]. The free coefficients of the functions can be fitted to the data using a fitting technique such as linear least squares regression (LLSR) [Manke et al., 2002; Low et al., 2005; King et al., 2009a,b, 2011; Rijkhorst et al., 2011]. In particular in the case where there are many highly correlated surrogates it may be beneficial to use a statistical technique for fitting which reduces such colinearities. Examples of such techniques used to establish direct correspondences are canonical correlation analysis (CCA) [Gao et al., 2008], partial least squares regression (PLSR) [Ablitt et al., 2004] and PCA [Zhang et al., 2010; Manke et al., 2003; Fayad et al., 2010].

In contrast to direct correspondences, some motion models establish an *indirect* correspondence between the surrogate and the motion data via a number of internal variables. That is, to apply the motion model the internal variables are optimised to find the best match between the measured surrogate data and the estimates of the surrogate data made by the motion model [White et al., 2009; King et al., 2012].

Some motion models show characteristics of both indirect and direct correspondence models [Peressutti et al., 2012].

As was discussed in Section 2.3 respiratory motion may exhibit significant variations from breathing cycle to breathing cycle (inter-cycle variability) and within breathing cycles (intra-cycle variability). The ability to capture these variabilities depends on two factors. First, the surrogate data need to be able to reflect these variabilities. Secondly, the data used to train the model must include such variabilities in respiratory motion. Note that for many applications, modelling no, or only a part of the respiratory motion variabilities may be sufficient. However, with the increasing resolution of many imaging modalities, such as PET, accounting for them could result in improvements in image quality.

Models that cannot account for any variabilities are typically simple relationships between a single respiratory trace and the motion descriptors [Wang et al., 1995; Manke et al., 2002; Rijkhorst et al., 2011]. Ablitt et al. [2004] proposed a motion model that in theory

Paper	Application	Surrogate	Motion Description	Model
Wang et al. [1995]	Cardiac MR acq.	From HD	Points of interest	Linear (fixed)
Manke et al. [2002]	Cardiac MR acq.	PB. Nav. (1)	Affine mat. coeffs.	Linear (LLSR)
Ablitt et al. [2004]	General cardiac MR	Chest surf. mesh (HD)	B-spline grid	Linear (PLSR)
Rijkhorst et al. [2011]	US-g. HIFU interv.	From HD (1)	B-spline grid	2nd ord. polynom. (LLSR)
Blackall et al. [2006]	Tumour mot. quant.	From HD (1)	Surface mesh	3rd ord. polynom. (LLSR)
Low et al. [2005]	RT interv.	Spirometer (2)	Points of interest	Linear (LLSR)
Zhang et al. [2010]	CBCT acq.	From HD (2)	Def. field	Linear (PCA)
McClelland et al. [2006]	4D CT acq.	IR marker track. (1)	B-spline grid	B-spline (LLSR)
King et al. [2009a]	Xray-g. cardiac interv.	From HD (1)	Affine transf. params.	2nd ord. polynom. (LLSR)
White et al. [2009]	Liver MR acq.	k-space lines (HD)	Def. field	Indirect
Fayad et al. [2010]	PET/CT acq.	Chest surf. mesh (HD)	Def. field	Linear (PCA)
King et al. [2011]	PET/MR acq.	PB. Nav. (1)	B-spline grid	2nd ord. polynom. (LLSR)
Manke et al. [2003]	Cardiac MR acq.	PB. Nav. (3)	Affine mat. coeffs.	Linear (PCA)
Gao et al. [2008]	RT interv.	Chest surf. mesh (HD)	B-spline grid	Linear (CCA)
King et al. [2009b]	Xray-g. cardiac interv.	From HD (1)	Affine transf. params.	2nd ord. polynom. (LLSR)
Peressutti et al. [2013]	US-g. cardiac interv.	US images (HD)	Affine transf. params.	Indirect
Schweikard et al. [2000]	RT interv.	IR marker track. (1)	Points of interest	Linear (LLSR)
Isaksson et al. [2005]	RT interv.	Skin. markers (1)	Points of interest	Linear (Neural networks)
King et al. [2012]	PET/MR acq.	2D MR images (HD)	B-spline grid	Indirect

Table 3.1: Overview of the subject-specific motion models reviewed in this thesis. The column “Surrogate” indicates the type of surrogate and its dimensionality in parentheses. “HD” refers to more complex data typically of higher dimensionality, i.e. a dimensionality greater than 3. If a low-dimensional surrogate has been derived from higher-dimensional data the term “From HD” is used. The last column contains the type of model, and the method used to find the correspondence in parentheses. “fixed” means that no fitting was performed, but a fixed relationship between the motion and surrogate data was used.

Paper	Intra-Cycle Var.	Inter-Cycle Var.	Adaptivity
Wang et al. [1995]	×	×	×
Manke et al. [2002]	×	×	×
Ablitt et al. [2004]	×	×	×
Rijkhorst et al. [2011]	×	×	×
Blackall et al. [2006]	✓	×	×
Low et al. [2005]	✓	×	×
Zhang et al. [2010]	✓	×	×
McClelland et al. [2006]	✓	×	×
King et al. [2009a]	✓	×	×
White et al. [2009]	✓	×	×
Fayad et al. [2010]	✓	×	×
King et al. [2011]	✓	×	×
Manke et al. [2003]	✓	✓	×
Gao et al. [2008]	✓	✓	×
King et al. [2009b]	✓	✓	×
Peressutti et al. [2013]	✓	✓	×
Schweikard et al. [2000]	×	×	✓
Isaksson et al. [2005]	✓	×	✓
King et al. [2012]	✓	✓	✓

Table 3.2: Ability to capture and adapt to respiratory motion variabilities of the reviewed subject-specific motion models.

could account for intra- as well as inter-cycle variabilities. The authors used a number of chest and abdominal surface points as the surrogate. However, the model was built from data which was acquired in a gated fashion based on a one-dimensional MR pencil beam navigator and without separating inspiration from expiration. The data acquired in this manner represents an average breathing cycle and does not capture any variability.

Models that are able to capture only intra-cycle variability typically use only a single one-dimensional surrogate, but in some way distinguish between inspiration and expiration. This can be, for example, by using the respiratory phase instead of an absolute amplitude value as a surrogate signal [McClelland et al., 2006], by additionally including a time-delayed version of the surrogate signal [Zhang et al., 2010; Isaksson et al., 2005] or by establishing separate correspondence models for inspiration and expiration [Blackall et al., 2006; King et al., 2009a,b, 2011]. A number of works in this category could potentially model inter-cycle variabilities as well, but as above, are limited by the type of the data. This applies in particular to motion models which are calibrated using 4D CT data such as Low et al. [2005] and Fayad et al. [2010]. 4D CT acquisitions are normally binned by respiratory phase but are averaged over multiple respiratory cycles and hence do not contain any inter-cycle variabilities.

In order to model inter-cycle variability, multiple or high-dimensional surrogate signals are required. Using multiple signals will generally also allow modelling of intra-cycle

variabilities. Therefore, models in this category can capture both types of variabilities. An early example of this is the previously mentioned work by Manke et al. [2003] who used up to three different MR pencil beam navigators and linearly related them to affine transformations of the heart. More recent approaches that are able to capture inter- and intra-cycle variabilities typically use high-dimensional data as the surrogate such as chest surface meshes [Gao et al., 2008] or high-dimensional imaging data [Peressutti et al., 2013; King et al., 2012]. An interesting special case is King et al. [2009b], who only derive a one-dimensional surrogate signal from imaging data but train the model on a number of different breathing types. In a classification step it is then decided which model is most appropriate at a given point in the application phase.

The ability to capture inter- and intra-cycle variabilities of the reviewed motion models is summarised in Table 3.2. Note that the above categorisation is based on the model's theoretical ability to capture said variations. To the best of my knowledge, to date there has been very little research into how well different kinds of surrogate data measure them.

3.3.3 Model Application

Lastly, for the duration of the treatment only the surrogate data are continually acquired and motion estimates are derived by using the surrogate data as input to the correspondence model. The motion estimates can be employed to correct for the motion with a temporal resolution determined by the frequency of the surrogate data acquisition.

For the majority of motion models the same surrogate data is available for the model formation as for the model application. In that case the model can simply be applied with the current surrogate value to derive the current motion estimate [Manke et al., 2002, 2003; Ablitt et al., 2004; Low et al., 2005; Zhang et al., 2010; McClelland et al., 2006; Gao et al., 2008; White et al., 2009; Fayad et al., 2010; Schweikard et al., 2000; King et al., 2011; Isaksson et al., 2005]. If this is not the case the two types of surrogate data must be in some way relatable to each other. For motion models which use a different imaging modality for the model formation than for the treatment guidance most often the diaphragm displacement is measured, which should lead to similar results independent of the modality as long as the diaphragm is visible [King et al., 2009a,b; Rijkhorst et al., 2011]. In a few instances, the surrogate data is of a different type all together, but is a subset of the training data. This is, for example, the case for King et al. [2012], where the model was formed using 3D MR dynamics and applied using 2D MR dynamics.

An underlying assumption of the majority of motion models is that the nature of the relationship between the surrogate data and the motion (i.e. the correspondence model) remains constant. For long treatment durations it is possible for the breathing motion to undergo significant inter-cycle variabilities, for example due to varying degrees of relaxation

of the patient during the procedure, because of pain or discomfort experienced [King et al., 2012; Hoogeman et al., 2009] or because of organ drift [Arnold et al., 2011; Von Siebenthal et al., 2007a]. In the traditional motion model paradigm the model is formed before the treatment and has no ability to adapt to changing breathing patterns.

A partial solution to this problem is to calibrate the model on data acquired during various anticipated breathing modes such as deep or shallow breathing [King et al., 2009b; Peressutti et al., 2013]. However, these models are not truly adaptive as they can only adjust to breathing patterns observed during the calibration phase. Furthermore, it may be impractical to acquire the necessary data for some patients. In scenarios where both the surrogate data and the calibration imaging data are also available in the application phase, it is possible to re-calibrate the model if necessary, allowing for *adaptive* motion modelling techniques. Table 3.2 also indicates which published techniques (discussed below) feature such adaptivity.

Most work on adaptive motion models has been done in the field of image-guided radiotherapy where X-ray images can be periodically acquired during the treatment to determine and verify the tumour location. Schweikard et al. [2000] proposed a motion model-based tumour tracking method where initially the motion of implanted tumour markers was correlated to surrogate data obtained from external chest markers. This initial model was then updated by intermittently acquired additional stereoscopic X-ray images along with the corresponding surrogate data. Each time a new tumour location was obtained in this way the oldest information was discarded and the model was recalibrated including the new tumour location and surrogate data. This solution is commercially available in the CyberKnife system². However, in the CyberKnife system new X-ray data can only be acquired every 1-6 minutes [Hoogeman et al., 2009; Seppenwoolde et al., 2007], which means the system can only adapt very slowly. In some clinics the accuracy of the system is assessed each time a new X-ray image is acquired. If the accuracy goes below a certain threshold all previously acquired data is discarded and the system is completely re-calibrated [Hoogeman et al., 2009]. In a simulation study, Seppenwoolde et al. [2007] investigated how the update frequency affects the tracking errors and found that faster updates (up to every 5 seconds) would lead to improved accuracy, however this would come at the cost of much higher radiation exposure. Another motion model for image-guided radiotherapy capable of adaptivity was proposed by Isaksson et al. [2005]. The authors employed an adaptive filter based on neural networks to predict the tumour location based on surrogate data. Additionally every 1-5 seconds new X-ray image and surrogate data were added to the model, which were used to update the correspondence model by automatically adjusting the filter parameters.

²<http://www.cyberknife.com/>

Two other scenarios where imaging and calibration data are available in the calibration phase as well as during the treatment are MR-guided treatments and PET/MR. King et al. [2012] proposed a scheme for PET/MR imaging which is calibrated using 3D MR data, but applied using fast 2D MR slices as surrogate data. Each time the model was applied an accuracy measure was evaluated, which was used to determine whether the model was still applicable. If not, the model was rebuilt by acquiring new 3D MR data. To the best of my knowledge, apart from that work, the potential of adaptive MR-guided methods remains unexplored.

3.4 Manifold Learning

Medical imaging datasets are typically represented, and processed (or visualised) in the very high-dimensional image space where each pixel/voxel adds an additional dimension. However, groups of images may vary due to many fewer degrees of freedom, and can be viewed as a set of points on a manifold of many fewer dimensions, embedded in the high-dimensional space. Linear and non-linear dimensionality reduction techniques can be used to uncover this low-dimensional structure. In the following the basic concept of dimensionality reduction and the most commonly used techniques will be briefly discussed.

This Section and the following are devoted to related work in the fields of manifold learning and manifold alignment, which are an essential part of the methodology chapters (Chapters 5-7) of this thesis.

3.4.1 Introduction

Linear techniques such as PCA [Wold et al., 1987] are widely used and have been applied to many applications in and outside of medical imaging. A recent example is the gating of PET data using a respiratory trace which was extracted from the list mode data using PCA [Thielemans et al., 2011]. In recent years, non-linear dimensionality reduction techniques, also collectively called *manifold learning* techniques, have gained much attention as a powerful tool to reduce the dimensionality of datasets which lie on a low-dimensional, non-linear manifold and a large number of manifold learning algorithms have been proposed in the last 15 years [Van der Maaten et al., 2009]. The existing techniques can be broadly divided into two categories: Techniques which preserve local properties of the data such as locally linear embedding (LLE) [Roweis and Saul, 2000], Hessian locally linear embedding (HLLE) [Donoho and Grimes, 2003], and Laplacian Eigenmaps (LEM) [Belkin and Niyogi, 2003], and techniques which preserve global properties of the data such as Isomap [Tenenbaum et al., 2000] and Diffusion Maps [Lafon and Lee, 2006]. Going into the details of those techniques is beyond the scope of this thesis, however, a brief introduction to LLE will be

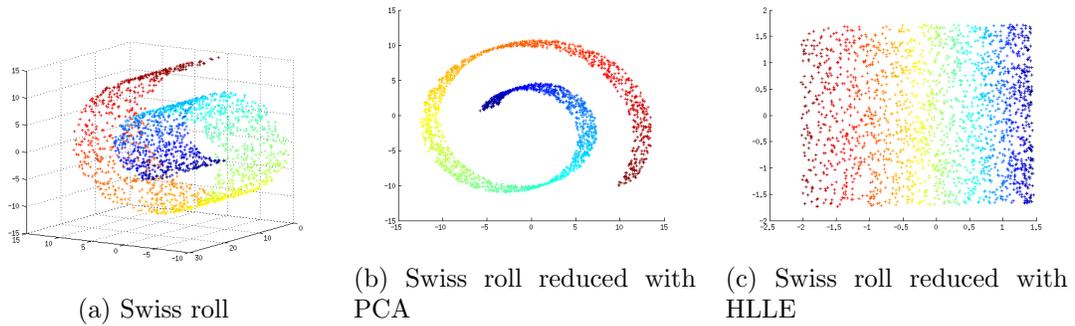


Figure 3.3: Manifold learning on the ‘Swiss roll’ manifold: (a) shows 2000 points sampled randomly on the Swiss roll manifold in 3D. (b) shows the reduction to 2 dimensions using PCA, (c) shows the same reduction using a manifold learning technique (HLLC).

given in Chapter 4.

The Swiss roll is an instructive, and much-cited example demonstrating the merits of manifold learning. Figure 3.3a shows 2000 points sampled randomly from the Swiss roll function. The data lies in 3-dimensional space, but is shaped like a rolled up piece of paper and thus inherently lies in 2-dimensional space. Uncovering the true shape using PCA is not feasible, as dimensionality reduction to two dimensions using PCA simply projects data onto a plane (Fig. 3.3b). Manifold learning, on the other hand, succeeds in correctly ‘unrolling’ the dataset (Fig. 3.3c).

Apart from their use in uncovering non-linear structure in high-dimensional data, manifold learning algorithms often have a very intuitive cost function and graph representation, which motivated many straightforward extensions. For instance, Lewandowski et al. [2010] suggested a simple extension of the LEM algorithm which incorporated temporal constraints into the cost function, i.e. high-dimensional points close in time should be embedded closely on the manifold. Another extension is *manifold alignment* which will be discussed in Section 3.5.

3.4.2 Manifold Learning For Understanding Physiological Motion

Time series of medical imaging data are often inherently high-dimensional. Consider, for example, a group of 2D MR images such as the one shown in Fig. 3.1b, which vary only due to deformations caused by respiratory motion. The dimensionality of these images equals the number of pixels in the images³. However, the underlying cause of the deformations can be captured in many fewer dimensions.

In recent years manifold learning was shown to be useful in the analysis of motion in such data, either directly on image intensities [Wachinger et al., 2011; Fischer et al., 2014]

³In this case 200×200 .

or on motion fields [Souvenir et al., 2006], making use of the fact that points which are close together in the low-dimensional space correspond to similar motion states. Applications include the extraction of respiratory traces for the gating of MR images [Wachinger et al., 2011] and from X-ray fluoroscopy images for motion modelling in image guided minimally invasive surgeries [Fischer et al., 2014], for the parametrisation of cardiac motion fields [Souvenir et al., 2006], and for the segmentation of cardiac CINE images [Zhang et al., 2006].

3.5 Manifold Alignment

Multiple related datasets often have a similar underlying low-dimensional manifold structure, but may not be directly comparable in high-dimensional space. Imagine some data sampled from two different Swiss roll manifolds (such as the one shown in Fig. 3.3) which are arbitrarily rotated in 3-dimensional space relative to each other. How could one go about finding corresponding points in each copy of the Swiss roll? Another example could be two sets of 2D MR images such as shown in Fig. 3.1b varying due to respiratory motion, but acquired at two different slice positions visualising different anatomy. In that case one may be interested in which slices were acquired at the same respiratory position. An elegant solution to establishing correspondences between such datasets is to reduce their dimensionality into a common low-dimensional embedded space and find the correspondences there. Such techniques are collectively known as manifold alignment techniques, and a number of them will be reviewed in the following. An overview of the examined techniques is given in Table 3.3.

Manifold alignment techniques have been applied to a variety of problems in computer vision and machine learning such as pose matching [Ham et al., 2005; Zhai et al., 2010; Toriki et al., 2010], image matching [Pei et al., 2012], articulated shape matching [Mateus et al., 2007, 2008], and for finding correspondences in document collections in different languages [Wang and Mahadevan, 2008] or about different topics [Wang and Mahadevan, 2009].

Recently, manifold alignment has also gained popularity in the field of medical image analysis. Applications include 4D CT reconstruction [Georg et al., 2008], ultrasound gating [Wachinger et al., 2011], extraction of regionally varying information such as cardiac and respiratory traces in free-breathing cardiac CINE images [Bhatia et al., 2012a,b], for Alzheimer’s classification [Bhatia et al., 2012b; Guerrero et al., 2014], for deriving modality independent image descriptors for multi-modal image registration [Wachinger and Navab, 2010, 2012], for the matching of cortical surface meshes [Lombaert et al., 2011, 2013], and lastly for decoupling function from anatomy in functional MR images [Langs et al., 2011, 2014; Nenning et al., 2015].

Paper	Application	Align. Meth.	Supervision	ML Meth.	# Dims.	# Dats.
<i>Applications to Medical Imaging:</i>						
Georg et al. [2008]	4D CT reconstruction	Two-step	Semi-supervised	Isomap	1	14
Bhatia et al. [2012a,b]	Regio. im. var.	One-step	Supervised	LEM	≥ 2	≥ 100
Wachinger and Navab [2012, 2010]	Multi-modal reg.	Two-step	Unsupervised	LEM	1	2
Wachinger et al. [2011]	US gating	Two-step	Supervised	LEM	1	29
Lombaert et al. [2011]	Surface matching	Two-step	Unsupervised (comp.)	LEM	≥ 2	2
Lombaert et al. [2013]	Surface matching	Both	Unsupervised (comp.)	LEM	≥ 2	2
Guerrero et al. [2006]	Alzheimer's class.	One-step	Semi-supervised	LEM	≥ 2	2
Nenning et al. [2015]	fMRI registration	Both	Unsupervised (comp)	Diff. Maps	≥ 2	12
Langs et al. [2011, 2014]	fMRI atlas	Two-step	Unsupervised (comp)	Diff. Maps	≥ 2	7
<i>Other Applications:</i>						
Ham et al. [2005]	Pose match.	One-step	Supervised	LEM	≥ 2	3
Zhai et al. [2010]	Pose match.	One-step	Semi-supervised	LLE	≥ 2	2
Torki et al. [2010]	Pose match.	One-step	Unsupervised (comp.)	LEM	≥ 2	7
Wang and Mahadevan [2008]	Language match.	Two-step	Semi-supervised	LLE	≥ 2	2
Wang and Mahadevan [2009]	Align. Doc. Colls.	One-step	Unsupervised	LEM	≥ 2	2
Mateus et al. [2007]	Shape match.	Two-step	Unsupervised	LLE	≥ 2	2
Mateus et al. [2008]	Shape match.	Two-step	Unsupervised	LEM	≥ 2	2
Pei et al. [2012]	Image match.	Two-step	Unsupervised	Isomap	≥ 2	2

Table 3.3: Overview of the manifold alignment techniques reviewed in this thesis. The table is divided into methods which were applied to problems in medical imaging and methods which were first proposed in other fields such as computer vision. Note that unsupervised techniques are further divided into techniques which require the data to be comparable in high-dimensional space (i.e. “comp”), and these which do not.

Some papers proposed techniques for aligning only two datasets [Lombaert et al., 2013, 2011; Guerrero et al., 2014; Zhai et al., 2010; Wang and Mahadevan, 2008, 2009; Wachinger and Navab, 2010, 2012; Pei et al., 2012; Mateus et al., 2008, 2007]. However, most methods can be easily extended to a larger number of datasets. In particular, in some medical imaging applications, naturally large numbers of datasets occur such as, for example, data from different image patches [Bhatia et al., 2012b,a], data from different US probe locations [Wachinger et al., 2011], or data from different subjects [Nenning et al., 2015].

3.5.1 Methods for Alignment

Irrespective of the number of datasets involved there are two general approaches to manifold alignment: 1) The multiple datasets are embedded separately using manifold alignment and then aligned in a subsequent alignment step using some shape matching technique. These approaches are known as *two-step* approaches. 2) Alternatively, the embedding and alignment can be written as one large optimisation problem, which can be solved in a single step. These approaches are known as *one-step* approaches. A small number of papers used a two-step approach to find a number of initial correspondences and then refined the findings using a one-step approach [Lombaert et al., 2013; Nenning et al., 2015]. The strategies used for each of the reviewed papers are summarised in the column “Align. Meth.” in Table 3.3.

Both one-step and two-step approaches are typically extensions of existing manifold learning techniques. The most commonly extended method is LEM, which has nice theoretical properties and is very easy to implement. Other techniques extended include LLE, Diffusion Maps and Isomap (see Table 3.3).

Two-step alignment methods employ various shape matching techniques to embeddings obtained using manifold learning. Methods used in the literature include simple re-scaling of the embeddings [Wachinger and Navab, 2012, 2010; Georg et al., 2008], affine transformation [Wachinger et al., 2011], orthogonal transformation [Mateus et al., 2008, 2007], Procrustes analysis [Wang and Mahadevan, 2008], and point cloud registration [Lombaert et al., 2011, 2013; Nenning et al., 2015]. The majority of two-step alignment methods are limited to linear transformations between the data. In order to account for more complex differences between the manifold embeddings a suitable non-rigid point cloud registration technique must be chosen as was done in Lombaert et al. [2011, 2013]. Moreover, as will be discussed in more detail in Chapter 4, each component of a manifold learning embedding can have an arbitrary sign, and the components’ ordering may vary between two separate embeddings. Two-step approaches must address these issues in some way [Mateus et al., 2008].

For one-step approaches the cost function of the employed manifold learning algorithm is extended by a term connecting the datasets. This leads to a single cost function encompassing all datasets. Often this expression can then be reformulated as an augmented version

of the original algorithm [Bhatia et al., 2012a; Ham et al., 2005; Zhai et al., 2010; Toriki et al., 2010; Wang and Mahadevan, 2009]. In contrast to two-step approaches, one-step approaches do not suffer from the problem of arbitrary ordering and sign of the embeddings' components. Furthermore, manifold embeddings which would vary non-rigidly from each other when embedded separately, can be naturally aligned, without the need for complex non-rigid methods.

3.5.2 Types of Supervision

Manifold Alignment can be performed in a supervised, semi-supervised or unsupervised manner.

In supervised techniques a one-to-one correspondence between all data points is already available prior to running the algorithm. One example is the simultaneous ultrasound gating and compounding method proposed by Wachinger et al. [2011]. The authors acquired free breathing abdominal data using an US wobbler probe which automatically sweeps a range of angles. The objective was to combine the data from all angles into a large compounded field of view. However, data acquired at each angle covered a range of respiratory positions such that the data had to be gated first, before the compounding could be performed. Due to the high acquisition frequency of ultrasound, the data acquired from all angles during a single sweep can be assumed to be at the same respiratory position. These correspondences were then used to affinely register the respiratory traces obtained using Laplacian Eigenmaps to obtain a consistent space that could be used for gating.

Semi-supervised techniques only have a subset of the one-to-one correspondences between datasets prior to running the algorithm. For example, Guerrero et al. [2014] aligned the manifolds of two different MR brain image datasets acquired from Alzheimer's patients to obtain a joint feature space for classification. One dataset was acquired using a 1.5T and the other using 3T MR scanner. For some patients both 1.5T and 3T scans were available, which provided the prior correspondences needed for alignment.

Lastly, unsupervised techniques require no prior correspondences. However, a large fraction of the methods in this category does require the data to be at least approximately comparable in the high-dimensional space such that initial correspondences can be established. For example, Lombaert et al. [2011] proposed a two-step manifold alignment technique for the alignment of cortical surfaces. The method was based on a cloud registration technique of embeddings obtained using Laplacian Eigenmaps. The registration step was constrained by the similarity of cortical features such as sulcal depth, cortical thickness and curvature, which were derived from each dataset prior to alignment. Similarly, Langs et al. [2011, 2014] initialised their method by comparing inter-subject voxel-wise fMRI signals, and Nenning et al. [2015] used spatial correspondences between voxels of different subjects

as a prior.

A few techniques are completely unsupervised and require neither prior correspondences nor comparability in the high-dimensional space. Wachinger and Navab [2012, 2010] proposed embedding patches of images acquired using different modalities. By aligning these two embeddings a modality-independent image descriptor could be obtained. However, the method was only demonstrated with one-dimensional embeddings, where the problem of alignment amounts to a simple rescaling of the embeddings.

A number of authors have proposed completely unsupervised methods in the field of computer vision.

Mateus et al. [2007] proposed a two-step method for the matching of three-dimensional humanoid shapes in different positions. The authors estimated an orthogonal transformation between the two point clouds using an expectation maximisation algorithm. The method was later extended to use Laplacian Eigenmaps [Mateus et al., 2008].

Another completely unsupervised two-step approach was proposed by Pei et al. [2012]. For each dataset the authors calculated the distances of each data point to its k -nearest neighbours. Next, for each point the distances to its neighbours were parametrised using a B-spline function, which were then compared across datasets to establish a set of initial correspondences. A non-linear optimisation framework was then used to non-rigidly align the manifolds.

The only completely unsupervised one-step manifold alignment method in the literature was proposed Wang and Mahadevan [2009]. The authors established initial correspondences between two datasets by considering similarities of the local neighbourhood within each of the datasets and subsequently aligned them in a one-step approach. Unfortunately, this approach required iterating through all of the permutations of the local neighbourhood of each point, which is only computationally tractable for very small neighbourhoods.

3.6 Discussion

In this Chapter, a review of solutions to some of the problems posed by respiratory motion has been presented, and furthermore manifold learning and manifold alignment techniques and their applications to medical image analysis were examined. The aim of this thesis is to investigate the utility of manifold alignment to provide improved solutions in some of the issues discussed above.

MR is a powerful modality allowing close to real-time acquisitions of high-contrast images with no radiation exposure. However, such imaging is currently limited to 2D MR slices. Slice stacking approaches have been proposed for a number of problems including PET/MR motion correction and accurate 4D MR imaging for radiotherapy planning and error quantification.

Most slice stacking techniques proposed for PET/MR (except the approach proposed by Dikaios et al. [2012]) rely on gating based on a one-dimensional respiration surrogate. Apart from the additional overhead that may be necessary to set up and evaluate such a signal, this limits the techniques' ability to capture intra- and inter-cycle variabilities. The resolution of currently used clinical PET scanners is around 5 mm [Bai and Brady, 2009]. However, the theoretical best resolution may be as low as 0.67 mm [Moses, 2011]. At such high resolutions inaccuracies due to respiratory variations are likely to become an issue, as they are already an issue in various other imaging techniques discussed above, such as MR-based coronary angiography.

A very accurate method for 4D MR reconstruction based on slice stacking was proposed by Von Siebenthal et al. [2007a]. However, as was discussed above the method requires a navigator slice to be acquired before each data slice, which is not used for the reconstructions. This significantly prolongs acquisition times. 2D MR slices are high-dimensional, but as was discussed in Section 3.4 often the underlying motion can be captured by fewer dimensions.

Manifold learning and alignment techniques have been successfully used to investigate physiological motion in medical images. As opposed to physically measuring motion e.g. using a MR pencil beam navigator, the dimensionality in which the motion should be examined can be set as a free parameter. Thus, manifold learning-based approaches have the potential to strike a balance between the unnecessarily high-dimensional navigator data used by Von Siebenthal et al. [2007a] and the simplistic surrogate measures usually used in gating. The feasibility of manifold alignment for slice stacking will be investigated in Chapter 5.

Another field where MR imaging is gaining importance is that of MR-guided interventions, in which MR can provide highly useful intra-procedure imaging. MR-guided HIFU treatments, as well as MR-guided cardiac interventions, have recently been performed in humans [Auboiroux et al., 2014; Tzifa et al., 2010], and an experimental MR-guided radiotherapy system has been proposed by Raaymakers et al. [2009] and evaluated on respiratory phantoms [Stam et al., 2012]. As real-time imaging is currently limited to a single plane, motion modelling approaches which can provide the full 3D motion may gain in importance. Motion modelling for MR-guided interventions has the potential to capture a significant amount of respiratory variabilities because sufficiently complex surrogate data can be easily acquired [King et al., 2012]. Moreover, adaptivity to previously unseen motion patterns can be implemented as the surrogate, as well as the imaging data, are available during the treatment. As was discussed in Section 3.3.3, the last possibility remains largely unexplored. An adaptive motion modelling approach based on manifold alignment will be investigated in Chapter 6.

A number of manifold alignment techniques have been proposed for the analysis of medical images, most of which were discussed in Section 3.5. However, the vast majority of techniques require either prior correspondences or the data to be comparable in the high-dimensional space. Many medical imaging datasets which were acquired during the same physiological motion but from different views, such as for example from different ultrasound probe angles [Wachinger et al., 2011], may lie on similar manifolds. However, the high-dimensional data may not always be easily comparable in high-dimensional space. A number of completely unsupervised manifold alignment methods, which can handle embeddings in arbitrary dimensions have been proposed in the field of computer vision. However, to the best of my knowledge no such methods have been proposed in the field of medical imaging. In Chapter 7, a completely unsupervised manifold alignment technique for matching the respiratory positions of multiple imaging datasets will be explored.

Chapter 4

Background and Theory

Contents

4.1	Manifold Learning on One Dataset	49
4.2	The Problem of Misaligned Embeddings	51
4.3	Manifold Alignment of Multiple Datasets	52
4.3.1	Two-Step Alignment	52
4.3.2	One-Step Alignment	53

In this chapter the basics of manifold learning and manifold alignment, on which subsequent chapters will build, are discussed and the notation is introduced. The common challenge for all methods presented in this thesis is finding correspondences in respiratory position between multiple datasets derived from images acquired under free breathing from different anatomical positions of a subject. This problem is approached by matching such high-dimensional imaging data by mapping them into a common low-dimensional space using manifold learning techniques and establishing correspondences by aligning those embeddings. The basic methodology for aligning the manifold embeddings of two high-dimensional datasets, which is reviewed in this chapter, is the same for all methods discussed in this thesis. In each of the subsequent chapters the following theory is extended to suit a particular application.

In order to find low-dimensional embeddings of single datasets the manifold learning technique LLE [Roweis and Saul, 2000] was used and a brief introduction to it will be given in the following section. Manifold embeddings derived from multiple related datasets which have been sampled from the same manifold are generally not aligned. This problem is discussed in Section 4.2. In Section 4.3 two strategies for aligning the manifolds of two datasets are discussed. For an overview of the mathematical notation used in this chapter and the remainder of this thesis refer to Table 4.1.

Table 4.1: List of frequently used mathematical notations.

Variable	Size	Description
p, q	1	Anatomical positions p and q . Various used to denote 2D slice positions (Chapters 5, 6, and 7) or ultrasound sound views from different probe locations (Chapter 7).
b_p^i	–	The i -th image acquired at the anatomical position p .
c_p^i	–	2D motion field derived by registering b_p^i to an exhale slice.
D	1	The dimensionality of the input data.
d	1	The dimensionality to which the data gets reduced in the manifold learning or manifold alignment step.
τ_p	1	Total number of data points in the dataset derived from anatomical position p .
L	1	The number of datasets to be aligned.
x_p^i	$D \times 1$	Vectorised image b_p^i (in Chapters 5 and 7), or vectorised motion field c_p^i (in Chapter 6).
X_p	$D \times \tau_p$	Matrix containing all τ_p data points x_p^i from anatomical position p .
y_p^i	$d \times 1$	Low-dimensional point corresponding to x_p^i .
Y_p	$d \times \tau_p$	Matrix containing τ_p low-dimensional points.
W_p	$\tau_p \times \tau_p$	Reconstruction weight matrix in LLE cost function (see Eq. (4.2)).
M_p	$\tau_p \times \tau_p$	Centred version of the weight matrix W_p .
U_{pq}	$\tau_p \times \tau_q$	Similarity kernel matrix connecting data from different anatomical positions p and q .

4.1 Manifold Learning on One Dataset

In natural datasets the dimensionality is often artificially high. Consider, for example, a set of τ_p coronal 2D MR slices $b_p^i, i \in \{1, \dots, \tau_p\}$ such as the one shown in Figure 3.1b, which differ from one another by a deformation of tissue due to respiratory motion. The vectorised images x_p^i lie in the very high-dimensional image space \mathbb{R}^D , where D is the number of pixels per image. However, the images vary due to a much smaller number of degrees of freedom. Hence, they lie on or near some smooth manifold \mathcal{M} of much smaller dimensionality $d \ll D$ which is embedded in the high-dimensional space. Manifold learning techniques map the high-dimensional coordinates of each of the data points x_p^i to the low-dimensional internal coordinates $y_p^i \in \mathbb{R}^d$ on the manifold. This reasoning applies not only to coronal MR slices but also to other types of imaging data, as well as to motion fields c_p^i derived from imaging data as will be discussed in Chapter 6. In the following, high-dimensional datasets containing all data points x_p^i acquired at an anatomical position p will be denoted by $X_p \in \mathbb{R}^{D \times \tau_p}$. Their low-dimensional counterpart will be referred to as

$Y_p \in \mathbb{R}^{d \times \tau_p}$.

LLE is a popular manifold learning technique due to its simplicity and intuitive cost function, which is based on elementary geometric intuitions. The algorithm starts off by first forming a k -nearest neighbour graph of the data based on the \mathcal{L}_2 -distance between the data points. The underlying idea is that each high-dimensional point x_p^i lies on an approximately linear patch together with its nearest neighbours. In order to characterise the local geometry of each patch, LLE expresses each data point as a linear combination of its neighbours. The reconstruction error is measured by the following cost function

$$E(W_p) = \sum_i |x_p^i - \sum_{j \in \eta(i)} W_p^{ij} x_p^j|^2. \quad (4.1)$$

Here $\eta(i)$ is the neighbourhood of the data point i as defined by the k -nearest neighbour graph. The optimal reconstruction weights W_p^{ij} can be calculated in closed form under the additional constraint that they add up to one. The weights minimising Eq. (4.1) are invariant to rotations and rescaling. The constraint additionally makes them invariant to translations. This allows the high-dimensional coordinates of each point x_p^i to be mapped into the global coordinates of the manifold by a linear mapping without changing the local geometry. A d -dimensional embedding ($d \ll D$) preserving this locally linear structure is given by the $Y_p \in \mathbb{R}^{d \times \tau_p}$ which minimise the following cost function:

$$\phi(Y_p) = \sum_i |y_p^i - \sum_{j \in \eta(i)} W_p^{ij} y_p^j|^2. \quad (4.2)$$

Eq. (4.2) reflects the fact that the low-dimensional coordinates y_p^i and their respective nearest neighbours are required to follow the same local geometry as the high-dimensional data in Eq. (4.1). This cost function can be rewritten in matrix form as

$$\phi(Y_p) = \text{Tr}(Y_p M_p Y_p^\top), \quad (4.3)$$

where $M_p = (I - W_p)^T (I - W_p)$ is the centred weight matrix, and $\text{Tr}(\cdot)$ is the trace operator. In order to make the problem well-posed the following two constraints need to be introduced:

$$\frac{1}{\tau_p} \sum_i (y_p^i)^\top y_p^i = I \quad (4.4)$$

$$\sum_i y_p^i = 0 \quad (4.5)$$

Eq. (4.4) requires the covariance of Y_p to be unity, which prevents degenerate embeddings. Eq. (4.5) requires the resulting embedding to be centred around the origin.

The cost function in Eq. (4.3) can be minimised subject to those constraints by calculating the Eigendecomposition of M_p . The embedding Y_p is given by the eigenvectors corresponding to the second smallest to $d + 1$ smallest eigenvalues of M . The eigenvector corresponding to the smallest eigenvalue represents the free translation component and discarding it fulfils Eq. (4.5) [Saul and Roweis, 2003]. However, the resulting embedding is not unique. As will be discussed in the following section, the embedding may vary due to flipping of each dimension because of the fact that the signs of the eigenvectors of M_p are arbitrary.

4.2 The Problem of Misaligned Embeddings

Separate datasets generated by the same mechanics, e.g. respiration, typically lie on similar manifolds because the datasets vary due to the same source of motion and hence have similar degrees of freedom. It is known that manifold embeddings¹ \tilde{Y}_p and \tilde{Y}_q derived from two datasets X_p and X_q acquired from two different anatomical positions p and q , but during the same physiological motion, have very similar shapes [Wachinger et al., 2011]. This knowledge can be used to identify corresponding data points in the two datasets.

Unfortunately, embeddings obtained from different datasets are generally not aligned in the low-dimensional space as they may vary due to flipping or rotations of the eigenvectors, and slight variations in the manifold structure. In Fig. 4.1 two such manifold embeddings into $d = 2$ dimensions derived from two high-dimensional datasets are shown. The datasets were derived from 2D MR slice data from two neighbouring slice positions which were acquired during a single scan, such that they can be expected to lie on very similar manifolds. As is illustrated in Fig. 4.1a the manifold embeddings may be flipped randomly in each dimension with respect to the other embedding. In this case the embeddings differ due to a flipping across the horizontal axis. In Fig. 4.1b the flipping has been manually accounted for by changing the sign of the vertical component of the red embedding. Even though the shapes now roughly match, it can be observed that the manifolds still vary due to a small rotation. This misalignment would prevent accurate matching of corresponding respiratory states in the two datasets. Clearly, simply flipping the sign of the eigenvectors is not enough to align the manifolds. In the following more advanced strategies for aligning manifold embeddings of two datasets are discussed.

¹I use the tilde to denote manifold embeddings which are not yet aligned to other manifold embeddings.

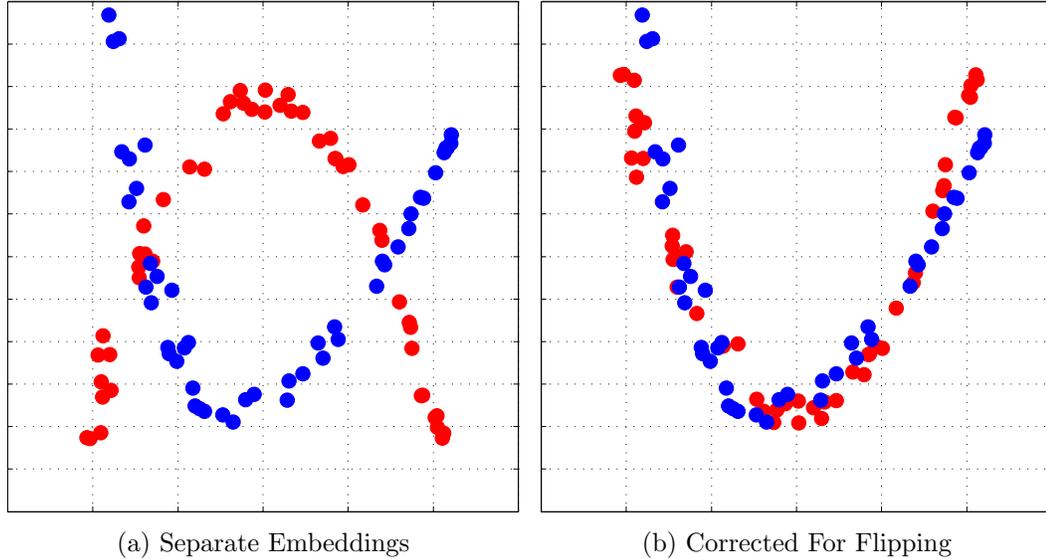


Figure 4.1: LLE embeddings obtained from two series of 50 coronal 2D MR slices acquired at two neighbouring slice positions p and q during the same MR scan. (a) shows the original embeddings \tilde{Y}_p, \tilde{Y}_q (red and blue). The red embedding is flipped with respect to the blue embedding due to the random direction of the eigenvector. In (b) the flipping has been manually corrected, however the embeddings still vary by a small but significant rotation.

4.3 Manifold Alignment of Multiple Datasets

As mentioned in Chapter 3 there are two general strategies for aligning the low-dimensional manifold embeddings originating from multiple high-dimensional datasets. The high-dimensional data can either be embedded in low-dimensional space separately and the resulting shapes can be matched in a subsequent alignment step (two-step approach), or the datasets can be embedded simultaneously in a one-step approach. Both these strategies are discussed in the following.

4.3.1 Two-Step Alignment

In order to align two datasets X_p and X_q , generated from data acquired from two different anatomical positions p and q , in a first step they can be reduced in dimensionality using LLE to obtain two unaligned low-dimensional embeddings \tilde{Y}_p and \tilde{Y}_q , such as shown in Fig. 4.1a.

One solution to aligning such shapes is to apply an algorithm designed to match two clouds of points, such as the iterative closest point (ICP) technique. In this thesis, the ICP implementation by Kroon [2011] was used, which can be applied to embeddings in two or three dimensions and can compute rigid or affine transformations. This algorithm,

however, cannot account for the random flipping in each dimension of the embeddings. To solve this problem, some corresponding points in both datasets have to be automatically identified. As is described in more detail in Section 5, for a series of coronal 2D MR slices exhale positions can be very reliably identified by looking at the mean intensity of the slices. The flipping can be accounted for by permuting through all possible 2^d (here $d = 2$ or $d = 3$) combinations and choosing the one which minimises the distance between the corresponding points. For other imaging datasets identifying corresponding exhale images may be more challenging. Next, one of the datasets is chosen as the reference and ICP is applied to obtain the aligned embeddings Y_p, Y_q .

I am not aware of other works which implement this approach for manifold alignment without *a priori* correspondences. However, the idea has been mentioned anecdotally by Lafon and Lee [2006], and Wachinger and Navab [2012].

4.3.2 One-Step Alignment

A different approach to finding aligned manifold embeddings Y_p, Y_q of two high-dimensional datasets is to embed them simultaneously. The cost function of LLE lends itself ideally to be extended to two datasets. The following is conceptually similar to the works by Ham et al. [2005] and Zhai et al. [2010] who, in contrast to this work, assumed the existence of *a priori* correspondences, and the extension of Laplacian Eigenmaps to multiple datasets by Toriki et al. [2010].

The problem of finding a simultaneous embedding can be written as a minimisation problem of the following cost function

$$C_{tot}(Y_p, Y_q) = \phi(Y_p) + \phi(Y_q) + \mu \cdot \psi(Y_p, Y_q), \quad (4.6)$$

where ϕ is the embedding error within the respective datasets p and q (intra-dataset cost functions) as given by Eq. (4.2), and ψ is the embedding error *between* the two datasets (inter-dataset cost function). This term ensures that corresponding points are embedded close to each other. Note that the inter-dataset cost is not known beforehand but is usually defined using some similarity between data points at runtime. The parameter μ regulates the influence of the inter-dataset cost function ψ on the embedding.

The cost function ψ can be defined as follows

$$\psi(Y_p, Y_q) = \sum_{i,j} |y_p^i - y_q^j|^2 U_{pq}^{ij}, \quad (4.7)$$

where

$$U_{pq}^{ij} = K(x_p^i, x_q^j),$$

is a (non-symmetric) similarity kernel relating data points from dataset p to q . This kernel is application specific and multiple choices were investigated in this thesis. Generally, the kernel was chosen to be of the form

$$K(x_p^i, x_q^j) = \exp\left(-\frac{\epsilon(x_p^i, x_q^j)^2}{2\sigma^2}\right), \quad (4.8)$$

where $\epsilon(\cdot, \cdot)$ is a distance function which must be defined such that the kernel $K(\cdot, \cdot)$ takes large values for similar data points and small values for dissimilar data points. The similarity kernel can be written as a matrix U_{pq} with high values connecting similar images from anatomical positions p and q . In Chapter 5, intensity-based distances between slices from neighbouring positions p and q are used to define the function ϵ . In Chapter 6 additionally a kernel based on motion similarities in the slice overlap is introduced for slices with different orientations. Lastly, in Chapter 7 a similarity kernel is introduced which is based solely on similarities in the local graph structure of each dataset in order to avoid comparisons in the high-dimensional space.

In Appendix A.1, I show that independent of the kernel choice the total cost function in Eq. (4.6) can be rewritten in matrix form as

$$C_{tot}(Y_p, Y_q) = Tr\left(\left[\begin{array}{c} Y_p^\top \\ Y_q^\top \end{array}\right]^\top \left[\begin{array}{cc} M_p + \mu D_{pq} & -\mu U_{pq} \\ -\mu U_{pq}^\top & M_q + \mu D_{qp} \end{array}\right] \left[\begin{array}{c} Y_p^\top \\ Y_q^\top \end{array}\right]\right), \quad (4.9)$$

where M_p, M_q are the respective centred weight matrices and D_{pq}, D_{qp} are diagonal matrices containing the column sums and row sums of U_{pq} , respectively. This problem now has the same form as the standard LLE cost function in Eq. 4.3, that is,

$$C_{tot}(Y_p, Y_q) = Tr(VHV^\top), \quad (4.10)$$

where H is the augmented matrix from Eq. (4.9) and V are the augmented embeddings. Under the constraints that

$$\sum_i y_p^i + \sum_i y_q^i = 0 \quad (4.11)$$

and

$$\frac{1}{\tau_p} \sum_i (y_p^i)^\top y_p^i + \frac{1}{\tau_q} \sum_i (y_q^i)^\top y_q^i = I \quad (4.12)$$

the simultaneous aligned embeddings are given by the second smallest to the $(d+1)$ -th smallest eigenvectors of L .

In the following chapters of this thesis different strategies for aligning the embeddings of not just two datasets, but large numbers up to around 40 are investigated. There, also

the choices for the similarity kernel $K(\cdot, \cdot)$ will be discussed which are needed to define the matrix U_{pq} .

Chapter 5

Simultaneous Groupwise Manifold Alignment

Contents

5.1	Introduction	56
5.2	Materials and Methods	58
5.2.1	Slice-by-Slice Acquisition	58
5.2.2	Manifold Alignment for Multiple Datasets	59
5.2.3	Choices for Inter-Dataset Kernel	62
5.2.4	4D Volume Reconstruction	67
5.3	Experiments	68
5.3.1	Compared Methods	68
5.3.2	Comparison of Manifold Alignment Methods	70
5.3.3	Comparison to State-of-the-Art Slice-Stacking Methods	72
5.4	Results	73
5.4.1	Comparison of Manifold Alignment Methods	73
5.4.2	Comparison to State-of-the-Art Slice-Stacking Methods	77
5.5	Discussion	79

5.1 Introduction

Current 3D MR imaging technology is limited by the lack of contrast in dynamic MR images of the lungs due to a rapidly decaying MR signal from the lung parenchyma. Additionally, dynamic 3D MR images suffer from low image resolution and relatively long acquisition times, which can lead to motion blurring, further limiting the accuracy of motion estimation.

Dynamic 2D MR scans can be acquired in a shorter time frame, have excellent in-plane resolution and high contrast in the lungs due to the in-flow of previously unexcited blood, but lack the coverage of 3D scans. Examples of 2D and 3D MR scans were shown in Fig. 3.1.

In order to make use of the good image properties of 2D MR, slice stacking approaches have been proposed in the literature. Dikaios et al. [2012] and Würslin et al. [2013] used slice stacking approaches for motion correction in PET/MR. Such approaches have the potential to very accurately correct for respiratory motion in the whole thorax including the lungs. The work by Würslin et al. [2013] relies on a 1D pencil-beam navigator to capture respiratory motion and thus cannot capture all respiratory variabilities. The method proposed by Dikaios et al. [2012] can potentially capture more of the respiratory variabilities, but, as will be shown in this chapter, lacks robustness.

Von Siebenthal et al. [2007a] used slice stacking to build very long 4D MR sequences in order to quantify respiratory variabilities and errors in radiotherapy [Von Siebenthal et al., 2007c], and to be used as input for a population-based motion model [Von Siebenthal et al., 2007b]. The approach was able to reconstruct very accurate 4D MR sequences capturing all intra- and inter-cycle respiratory variabilities, but was hampered by a low scan efficiency as only 50% of the acquired data were used for reconstruction.

The aim of this chapter is to investigate the utility of manifold alignment for creating a robust, navigator-less and 100% efficient slice stacking technique for retrospectively reconstructing 4D MR sequences from sequentially acquired 2D MR slices. Such a technique has potential use for the same applications as the method proposed by Von Siebenthal et al. [2007a]. It can also be used to motion correct PET data in a hybrid PET/MR acquisition as will be demonstrated.

In the following section, a number of different ways to extend the manifold alignment theory from Chapter 4 to large numbers of datasets consisting of MR slices acquired at different slice positions are examined. In the Experiment and Results sections (i.e. Sections 5.3 and 5.4) the manifold alignment methods are evaluated for the 4D MR reconstruction from real and synthetic slice-by-slice MR data. The evaluations include a comparison between the proposed extensions of manifold alignment and another manifold alignment technique from the literature (i.e. Toriki et al. [2010]), and a comparison against the slice stacking methods proposed by Würslin et al. [2013] and Dikaios et al. [2012]. Furthermore, in Section 5.3.3 it will be demonstrated how a manifold alignment-based slice stacking method can be used to correct PET/MR data for respiratory motion using synthetic PET/MR data derived from real 4D CT scans. Lastly, Section 5.5 contains the discussion.

The majority of the work in this chapter has been presented in Baumgartner et al. [2013] and Baumgartner et al. [2014a].

5.2 Materials and Methods

The methods proposed in this chapter for the reconstruction of 4D MR sequences from 2D MR slices follow a basic scheme shown in Fig. 5.1. In a first step, data is acquired in the slice-by-slice protocol outlined in Section 5.2.1. The slices acquired from each slice position \mathcal{C}_p form a dataset X_p . Next, respiratory correspondences between the datasets (i.e. data collected from the different slice positions) are established using a manifold alignment approach. The different approaches for extending the theory outlined in Chapter 4 are discussed in Sections 5.2.2 and 5.2.3. Lastly, the found correspondences are used to reconstruct a 3D MR volume for each 2D MR input slice, resulting in a 4D MR sequence. This step is described in Section 5.2.4.

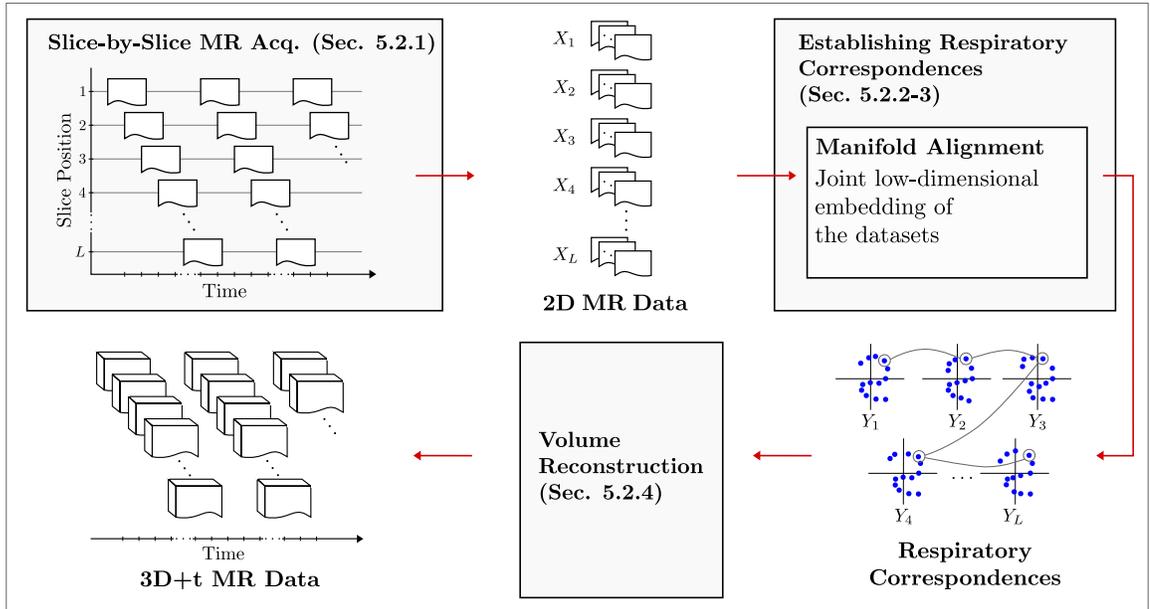


Figure 5.1: Schematic illustrating the 4D MR reconstruction from 2D MR data using manifold alignment. First, 2D MR data is acquired in a slice-by-slice fashion. Respiratory correspondences between different anatomical positions are established by means of manifold alignment of data from different slice positions. The X_p denote data acquired at slice position \mathcal{C}_p , the Y_p denote low-dimensional manifold embeddings thereof. Lastly, for each 2D input slice, a 3D volume is reconstructed based on the established correspondences.

5.2.1 Slice-by-Slice Acquisition

The volume of interest was imaged by sequentially acquiring coronal slices b_p^i at shifting slice positions p covering the whole region of interest, as illustrated in Figure 5.1. In order to sufficiently sample all respiratory states each slice position was acquired 50 times. The vectorised image b_p^i is denoted by x_p^i and all data collected from a slice position p make up

dataset X_p (see Table 4.1). The acquisition time for each slice was 160 ms. To maximise vessel contrast, and to minimise cardiac motion, only one slice was acquired per heartbeat at systole resulting in a temporal resolution of approximately one slice per second.

The acquisitions were carried out on a Philips Achieva 3T MR scanner using a T1-weighted gradient echo sequence with an acquired in-plane image resolution of $1.4 \times 1.4 \text{ mm}^2$, a slice thickness of 8 mm, repetition and echo times (TR and TE) of 3.1 and 1.9 ms, a flip angle (FA) of 30 deg, and a SENSE factor of 2. To cover a region of interest in the thorax including most of the lungs in the A-P direction typically 29-34 slice positions were needed, which resulted in a total acquisition time of 24-28 minutes. In order to decrease the effective slice thickness of the reconstructed volumes, and to help to establish respiratory correspondences between adjacent slice positions, the slice positions were defined with a 4 mm overlap. This resulted in an effective slice thickness of 4 mm in a reconstructed volume. To avoid polarisation artefacts from previous slices, a standard¹ slice position sampling scheme was used in which the slice position number is increased with a step size equal to the rounded square root of the total number of slice positions.

5.2.2 Manifold Alignment for Multiple Datasets

In the following, three methods are presented for extending the theory for manifold alignment of two datasets in Chapter 4 to L datasets: (1) An extension of the two-step approach, (2) An extension of the one-step approach, (3) a group-wise one-step approach.

Extension of Two-Step Approach

The extension of the two-step approach outlined in Section 4.3.1 for 2 datasets is straightforward and very similar to existing work in the literature such as Lombaert et al. [2011].

In a first step, the dimensionality of all datasets X_1, \dots, X_L is reduced using LLE to obtain unaligned embeddings $\tilde{Y}_1, \dots, \tilde{Y}_L$. The slice position $ref = \text{round}(L/2)$ is chosen as reference to which the others will be aligned to.

As discussed in Section 4.3.1, the components of the embeddings have arbitrary signs. Therefore, in a next step this “flipping” is accounted for. To this end a number of exhale images are identified in each dataset. In coronal images such as those acquired for this chapter all slice positions show the lungs which vary in size due to respiratory motion. Therefore the mean intensity \bar{m}_p^i of a data point x_p^i can give an approximate indication of its position in the respiratory cycle. The x_p^i with the maximum \bar{m}_p^i within each slice position is labelled as exhale. Then all embeddings are flipped to minimise the distance of their exhale point to the exhale point of \tilde{Y}_{ref} as described in Chapter 4.

¹This sampling scheme is called “interleaved” on the scanner model used to acquire the data.

Next, ICP is applied in a pair-wise fashion to align each of the embeddings \tilde{Y}_p to \tilde{Y}_{ref} . This produces the aligned embeddings Y_1, \dots, Y_L .

Extension of One-Step Approach

One-step approaches can be generalised from two datasets to an arbitrary number of datasets. To this end a joint cost function based on LLE is formulated and then a low-dimensional representation is obtained simultaneously for all datasets in a single embedding step. While this is a straight-forward generalisation of the theory in Chapter 4 for embedding only two datasets, to the best of my knowledge, no related works have extended LLE to embed $L > 2$ datasets simultaneously.

The extension is achieved by augmenting the cost function from Eq. (4.2) as follows:

$$C_{tot}(Y_1, \dots, Y_L) = \sum_{\ell=1}^L \phi_{\ell}(Y_{\ell}) + \frac{\mu}{2} \sum_{\substack{p=1, q=1 \\ p \neq q}}^L \psi(Y_p, Y_q), \quad (5.1)$$

where ϕ_{ℓ} is given by the LLE cost function defined in Eq. (4.2) and $\psi(Y_p, Y_q)$ is given by the usual inter-dataset cost function defined in Eq. (4.7). As in Chapter 4, μ is a weighting parameter which governs the influence of the inter-dataset terms. Analogous to the case with only two datasets in Chapter 4, C_{tot} can be rewritten in matrix form as $C_{tot} = Tr(VHV^T)$, where $V = [Y_1, \dots, Y_L]$ is the matrix containing the concatenated embeddings and

$$H = \begin{bmatrix} M^{(1)} + \mu \sum_p D_{1p} & -\mu U_{12} & \dots & -\mu U_{1L} \\ -\mu U_{21} & M_2 + \mu \sum_p D_{2p} & \dots & -\mu U_{2L} \\ \vdots & & \ddots & \vdots \\ -\mu U_{L1} & -\mu U_{L2} & \dots & M_L + \mu \sum_p D_{Lp} \end{bmatrix} \quad (5.2)$$

Here, again the diagonal degree matrices D_{pq} are given by $D_{pq}^i = \sum_j U_{pq}^{ij}$. Under the scaling and translation constraints

$$\sum_{\ell=1}^L v_{\ell}^T v_{\ell} = 1 \quad (5.3)$$

and

$$\sum_p \sum_i y_p^i = 0, \quad (5.4)$$

which are generalisations of the constraints in Eqs. (4.11) and (4.12), the embedding V is given by the second smallest to $d + 1$ smallest eigenvectors of H .

Simultaneous Groupwise Embedding

As shown above, the cost function in Eq. (4.9) from the Theory chapter (i.e. Chapter 4) for embedding two datasets simultaneously can be easily extended to L datasets. However, in that approach the inter-dataset similarity matrices U_{pq} must be defined for all combinations of p and q , even for slice positions which are not adjacent. Slice positions which are far apart may not be directly comparable in image space as they visualise different parts of the anatomy².

Therefore, here a novel approach is proposed which avoids this problem. The datasets are embedded simultaneously, in a one-step fashion, in overlapping groups of two. That is, L high-dimensional input datasets X_1, \dots, X_L are embedded in $L - 1$ groups,

$$\mathcal{G}(\mathcal{C}_1, \mathcal{C}_2), \mathcal{G}(\mathcal{C}_2, \mathcal{C}_3), \dots, \mathcal{G}(\mathcal{C}_{L-1}, \mathcal{C}_L),$$

where \mathcal{C}_p denotes the (coronal) slice position p^3 . Each group contains the embeddings Y_p, Y_{p+1} of two adjacent slice positions \mathcal{C}_p and \mathcal{C}_{p+1} . The groups are overlapping in the sense that data from each slice position appears in two neighbouring groups. For example, data from slice position 4, is embedded in group $\mathcal{G}(\mathcal{C}_3, \mathcal{C}_4)$ as well as group $\mathcal{G}(\mathcal{C}_4, \mathcal{C}_5)$. X_1 and X_N are embedded in only one group because there is no more data before or after them, respectively. Figure 5.2 shows an artificial example of a groupwise embedding in $d = 2$ dimensions and the relations between the groups. In the following, it will be explained how the groups can be connected.

Although the embeddings from each group could be transformed into a single space using some shape matching technique, this is in fact unnecessary. By embedding the datasets in overlapping groups in this manner, correspondences between the different datasets can be directly found by going from group to group. The two members of each group are aligned due to the simultaneous embedding, and the connections to the next group are deterministically known through the group overlap, i.e. because of the dataset that the groups share. For example, consider an arbitrary point Y_p^i (labelled with a square in Figure 5.2), which is embedded in $\mathcal{G}^{(p)}$, i.e. in $\mathcal{G}(\mathcal{C}_p, \mathcal{C}_{p+1})$. Since within the group the manifolds are aligned, the closest neighbour in Y_{p+1} can be found directly by evaluating the distances to Y_p^i in the low-dimensional space (see dotted lines in Figure 5.2). The nearest neighbour is indicated by a circle in group $\mathcal{G}(\mathcal{C}_p, \mathcal{C}_{p+1})$. The point Y_p^i and its newly found nearest neighbour Y_{p+1}^j can be looked up directly in their respective neighbouring groups

²Note that leaving the similarity kernel U_{pq} undefined for non-adjacent p and q was investigated in preliminary experiments and did not lead to acceptable results. The basic extension using a U_{pq} for all combinations of p and q is evaluated as the technique FULL in the experiments described in Section 5.3.

³In the following, the coronal slice positions will be referred to as \mathcal{C}_p instead of just p in order to be consistent with the notation used in Chapter 6, where additionally sagittal slices will be used in the groupwise embedding.

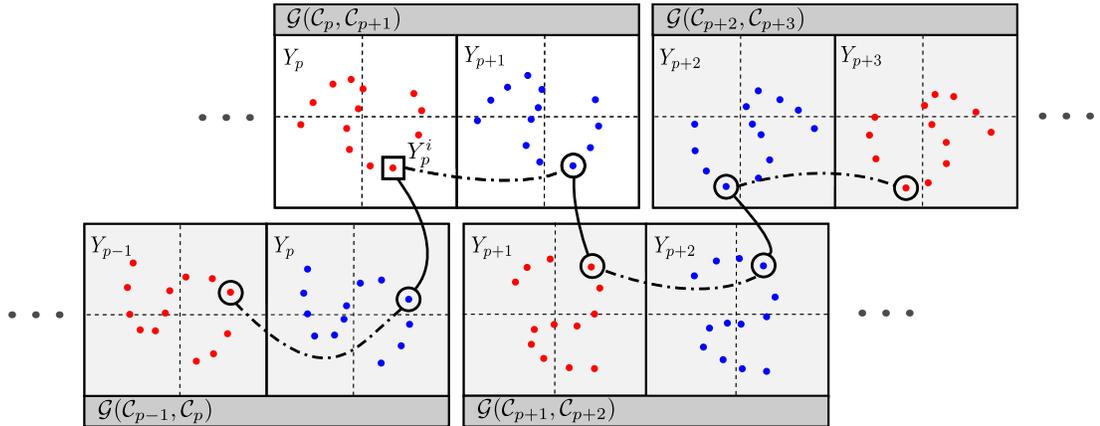


Figure 5.2: Schematic illustration of groupwise manifold alignment. The curved lines illustrate the connections between the manifold embeddings through the group overlap across groups (solid), or nearest neighbour look-up within aligned groups (dotted).

$\mathcal{G}(\mathcal{C}_{p-1}, \mathcal{C}_p)$, and $\mathcal{G}(\mathcal{C}_{p+1}, \mathcal{C}_{p+2})$, since those groups also contain a copy of Y_p , and Y_{p+1} , respectively. This look-up operation is denoted by the solid lines in Figure 5.2. From there, another nearest neighbour step can be performed, finding points on Y_{p-1} , and Y_{p+2} . In this manner the low-dimensional points closest to Y_p^i on all other manifold embeddings can be found iteratively.

Note that by choosing the closest neighbour within the groups (dotted lines) and transporting that between groups (solid lines) instead of the actual point coordinates, small introduced are incurred in the process of iteratively looking up correspondences. In order to propagate the actual point coordinates, interpolating the location of the point in the new group by looking at the k nearest neighbours on the current manifold embedding and finding an average of those points in the manifold embedding of the new group was investigated. However, in the present application of volume reconstructions, this only slightly influenced the results and therefore, for simplicity, only the closest neighbour is transported as described above.

5.2.3 Choices for Inter-Dataset Kernel

For the direct extension of the one-step approach and for the simultaneous groupwise manifold alignment (SGA) approach similarity kernels must be defined on the data. Two different approaches were investigated: 1) A similarity kernel based directly on image comparisons and 2) a registration-based similarity kernel additionally taking into account anatomical differences between slice positions.

Image Similarity-Based Inter-Dataset Kernel

An image-based inter-dataset kernel U_{pq} based directly on the image similarity of data from adjacent slice positions \mathcal{C}_p and \mathcal{C}_q can be defined as

$$U_{pq}^{ij} = K(x_p^i, x_q^j) = \exp\left(-\frac{\tilde{\mathcal{L}}_2(x_p^i, x_q^j)}{2\sigma^2}\right), \quad (5.5)$$

where $\tilde{\mathcal{L}}_2$ denotes the normalised \mathcal{L}_2 -distance, which is a scaled version of the \mathcal{L}_2 -distance such that the maximum $\tilde{\mathcal{L}}_2$ value between slices acquired at two adjacent slice positions is equal to 1. The parameter σ governs the kernel shape. The normalisation step allows the value of σ to be subject independent. The $\tilde{\mathcal{L}}_2$ -distance is chosen because the data is monomodal and because it gave superior results in preliminary experiments when compared to normalised mutual information and normalised cross correlation.

Registration-Based Inter-Dataset Kernel

Even though the overlapping slices used in this chapter cover some common anatomy and consequently look similar, the small changes in slice position between two adjacent slices still cause non-negligible differences in the images. In particular, a shift in slice position might cause an apparent deformation similar to a change in respiratory position (for example a shift in the diaphragm position, or a change in lung area). Thus the most similar looking slice may not always be the one with the closest respiratory position and basing the inter-dataset kernel directly on the image similarities may adversely affect the resulting reconstructions.

To address this, a novel registration-based inter-dataset kernel is proposed, which incorporates knowledge of the approximate relations between adjacent slice positions. Those relations are obtained using registration of exhale slices, which are subsequently transported to different respiratory states using transformations obtained from a second set of registrations. In this manner approximations of the relations between adjacent slice positions at any respiratory position are found, which can then be incorporated in the similarity kernel in order to improve the similarity measure between slice positions. This process is explained in more detail below. The relations between the slices and transformations are shown in Figure 5.3.

In a first step an exhale volume consisting of exhale slices $b_p^{(exh)}$ is reconstructed from the slice-by-slice data. Because the exhale state is very reproducible [Blackall et al., 2006] this can be accomplished by simply stacking the slices with the smallest lung area. As described in Section 5.2.2, exhale slices can be found reliably by calculating the mean intensity of each slice position; the slices with the highest mean intensity corresponded to the smallest lung area. Note that no additional imaging is necessary for this step.

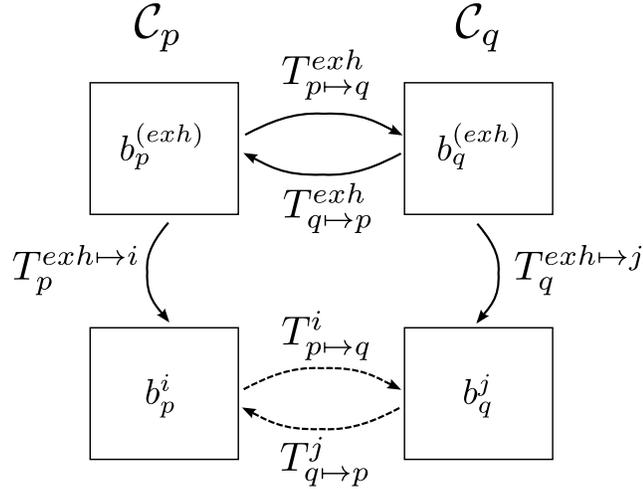


Figure 5.3: Relations between the slices at adjacent slice positions \mathcal{C}_p and \mathcal{C}_q and the different (2D) transformations between them. The transformations $T_{exh}^{p \rightarrow q}$, and $T_{exh}^{q \rightarrow p}$ between two exhale slices $b_p^{(exh)}$ and $b_q^{(exh)}$ were transported to different respiratory positions using the transformations $T_p^{exh \rightarrow i}$ and $T_q^{exh \rightarrow j}$, to arrive at approximate transformations $T_i^{p \rightarrow q}$ and $T_j^{q \rightarrow p}$ between two slices b_p^i and b_q^j at arbitrary time points i and j .

The relations between adjacent exhale slices at slice positions \mathcal{C}_p and \mathcal{C}_q are captured by the transformation $T_{p \rightarrow q}^{exh}$ which is obtained by registering $b_p^{(exh)}$ to its neighbour $b_q^{(exh)}$ using 2D to 2D B-spline registration [Rueckert et al., 1999] with a control point spacing of 2.8 mm, a smoothness penalty term of $\lambda = 0.01$ and sum of squared differences (SSD) as similarity measure. A relatively small control point spacing was chosen because changes between adjacent slice positions typically do not involve large displacements. The corresponding motion fields $D_{p \rightarrow q}^{exh}$ can be obtained by subtracting the grid locations (x', y') from the transformed grid points, i.e.

$$D_{p \rightarrow q}^{exh}(x', y') = T_{p \rightarrow q}^{exh}(x', y') - (x', y'), \quad (5.6)$$

where (x', y') are the grid points of the transformation.

In order to make this mapping across slice positions available at all respiratory positions, the transformations $T_{p \rightarrow q}^{exh}$ are transported to the different respiratory states. This can be viewed as transporting the 2D deformation fields from the exhale coordinate system to the coordinate systems of different respiratory states which are non-rigidly deformed with respect to the exhale coordinate system. Rao et al. [2002] proposed a framework to solve a conceptually similar problem, namely to transport deformation fields from one patient coordinate system to another. This framework is used here to solve the problem of transporting deformation fields to the coordinate systems of different respiratory states.

To obtain mappings to different respiratory states, additional 2D to 2D registrations to different time points within the slice positions are performed. That is, the transformations $T_p^{exh \rightarrow i}$ are obtained by registering $b_p^{(exh)}$ to b_p^i . For this step, B-spline registration with a larger control point spacing of 14 mm is used, because the deformations due to respiratory motion are larger than the deformations from slice position to slice position.

To obtain $T_{p \rightarrow q}^i$ the deformation fields $D_{p \rightarrow q}^{exh}$ are transported from the coordinate system (x', y') to the coordinate system of respiratory state i , which is denoted by (x, y) . If the motion vector at positional coordinates (x'_0, y'_0) in the coordinate system of exh is equal to $d_{p \rightarrow q}^{exh}$, the transported vector at the location (x_0, y_0) and respiratory position i is given by $d_{p \rightarrow q}^i$, where

$$(x_0, y_0) = T_p^{exh \rightarrow i}(x'_0, y'_0) \quad \text{and} \quad d_{p \rightarrow q}^i = J^{-1} \cdot d_{p \rightarrow q}^{exh}. \quad (5.7)$$

Here, J is the Jacobian matrix of $(T_{(p)}^{exh \rightarrow i})^{-1}$, and $(T)^{-1}$ denotes a numeric approximation of the inverse of T . The multiplication by J^{-1} is necessary to account for non-translational changes between the coordinate systems [Rao et al., 2002]. $T_{p \rightarrow q}^i$ can be obtained from the transported motion fields using Eq. (5.6). This process can be repeated going in reverse from slice position \mathcal{C}_q to \mathcal{C}_p to arrive at the transformation $T_{q \rightarrow p}^j$ at respiratory position j .

With approximate knowledge of the relations between the slice positions at arbitrary respiratory states an improved distance measure between the high-dimensional points x_p^i and x_q^j can be defined based on the corresponding images b_p^i and b_q^j as

$$\epsilon(x_p^i, x_q^j) = \frac{1}{2} \tilde{\mathcal{L}}_2(b_q^j, b_p^i \circ T_{p \rightarrow q}^i) + \frac{1}{2} \tilde{\mathcal{L}}_2(b_p^i, b_q^j \circ T_{q \rightarrow p}^j), \quad (5.8)$$

where, again $\tilde{\mathcal{L}}_2$ denotes the normalised \mathcal{L}_2 -distance, and $I \circ T$ is the transformation of image I with the transformation T . This means the similarities are evaluated after transforming the slice at \mathcal{C}_p into the coordinate system of \mathcal{C}_q at the corresponding respiratory position, and vice versa. Using both the registrations $T_{p \rightarrow q}^i$, and $T_{q \rightarrow p}^j$ improved the robustness of the similarity measure.

Note that the transformations $T_{p \rightarrow q}^i$ and $T_{q \rightarrow p}^j$ are only approximations, and might lose validity at deep inhale states where the slice relations change significantly with respect to an exhale state. However, even at those states the similarity measure is significantly improved over using the $\tilde{\mathcal{L}}_2$ -distance directly, as will be shown in the experiments.

The novel registration-based similarity kernel relating the slices at slice position \mathcal{C}_p to

the slices at slice position \mathcal{C}_q can then be defined as

$$U_{pq}^{ij} = K(x_p^i, x_q^j) = \exp\left(-\frac{\epsilon(x_p^i, x_q^j)^2}{2\sigma^2}\right), \quad (5.9)$$

where again σ governs the kernel shape.

Kernel Sparsification

Regardless of which similarity kernel is used, i.e. the image similarity based one in Eq. (5.5) or the registration-based one in Eq. (5.9), the quality of the embeddings benefits from a sparsification of the respective matrices U_{pq} . Note that in a simple k-nearest neighbour sparsification, neighbourhood relations are directed and thus not symmetric, i.e. if ‘ A is a neighbour of B ’ it does not necessarily follow that ‘ B is a neighbour of A ’. This means that the sparsification will depend on which slice position the k-nearest neighbour operation is based on, for example if the k-nearest connections of slice position \mathcal{C}_p to \mathcal{C}_{p+1} are kept a different result will be obtained than when looking for the k-nearest connections of slice position \mathcal{C}_{p+1} to \mathcal{C}_p . This in turn will lead to biased connectivities and will adversely affect the embedding. In practice, this means that some slices will be connected to many neighbours, while others will be connected to none. Furthermore, note that making the sparsified kernel symmetric by mirroring the connections would insert connections which do not correspond to any physical similarity. For example, if slice i on slice position \mathcal{C}_p is close to slice j on slice position \mathcal{C}_q (i.e. x_p^i is close to x_q^j) it does not follow that also x_p^j is close to x_q^i . x_p^j may in fact be completely different from x_q^i . In order to illustrate the adverse effects of using a nearest neighbour based sparsification, this method will also be evaluated in the experiment section (Section 5.3) as technique SGA.NN.

To overcome these limitations a kernel sparsification technique based on a global bipartite maximum edge similarity matching is used. That is, the matching is obtained in which every data point in X_p is connected to exactly one data point in X_q , and the sum of similarities over all of the corresponding edges U_{pq}^{ij} is maximised. The schematic in Figure 5.4 illustrates this process. The bipartite matching that maximises the similarity is highlighted in red in Figure 5.4b. This is equivalent to a combinatorial optimisation problem and can be solved using the Hungarian method [Kuhn, 1955]. Note that using this sparsification, if A is a neighbour of B , by definition B has to be a neighbour of A .

The resulting graph can be written as a sparse, symmetric matrix U_{pq} , which in every row and every column has exactly one non-zero entry, $0 < U_{pq}^{ij} \leq 1$. This is similar to the case of labelled connections as in Ham et al. [2005], with the difference that here the labels are not known *a priori*, but instead have a certainty measure (i.e. the kernel value) attached to them.

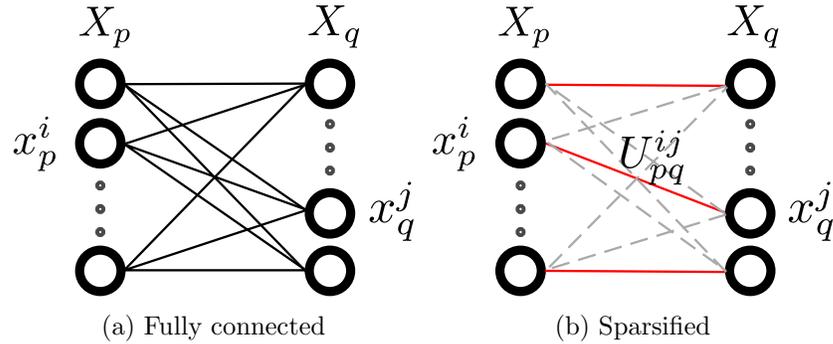


Figure 5.4: Sparsification of similarity kernel by one-to-one node matching. (a) shows the fully connected graph, (b) shows the optimal one-to-one mapping.

5.2.4 4D Volume Reconstruction

Computing aligned low-dimensional embeddings using any of the outlined approaches makes it possible to establish respiratory correspondences in the low-dimensional space. The next step is to reconstruct the slices into a sequence of 3D volumes. For a slice acquired from slice position \mathcal{C}_p , images from other slice position which are at the same respiratory position can be identified in the low-dimensional embedded space. Those slices can then be stacked into a coherent 3D MR volume. This process is performed for each slice x_p^i in the original acquisition order to arrive at a 4D MR (or 3D+t) reconstruction over the length of the acquisition. Note that because each slice position has been acquired 50 times (see Section 5.2.1) it can be assumed that most respiratory positions have been covered and there will be enough data to reconstruct a volume from each slice.

For the extensions of the two-step and one-step approaches, which were outlined in Section 5.2.2 and Section 5.2.2, respectively, the data from all slice positions lie in the same low-dimensional space. Thus, for a given input slice, it suffices to directly look up the nearest neighbours to it from all the other slice positions. The corresponding slices can then be stacked in a volume.

For the groupwise manifold alignment method the low-dimensional embeddings are not in a single consistent space, but rather in the spaces of the respective groups. However, the coordinates of the groups can be related to each other by making use of the group overlaps, as was explained in Section 5.2.2, in particular in Figure 5.2. In this manner, for each input slice the closest points in the low-dimensional space can be identified. Once identified the high-dimensional slices corresponding to the low-dimensional points can be stacked into volumes, as before.

5.3 Experiments

The experiments section consists of two parts answering the two research questions of this Chapter:

- Which of the outlined manifold alignment techniques performs the best for 4D MR reconstruction from 2D MR slice data?
- How do manifold alignment based methods compare against state-of-the art slice-stacking techniques from the literature?

The experiments were performed on synthetic (Experiments 1 & 3) and on real 2D MR slice-by-slice data (Experiments 2 & 4). Experiment 5 contains a demonstration of how a manifold alignment based method can be used in a PET/MR motion correction scenario. The experiments are summarised in Table 5.1 for the reader’s convenience.

Table 5.1: Summary of the experiments in this chapter including the data it was performed on, the methods that were evaluated and the evaluations which were performed.

Exp.	Data	Evaluated Techniques	Evaluations
1	synth. MR	MA techniques	Determination of the optimal μ parameter for all methods; Reconstruction error; Visual inspection
2	real MR	MA techniques	Reconstruction consistency; Visual inspection
3	synth. MR	State-of-the-art techniques	Reconstruction error; Visual inspection
4	real MR	State-of-the-art techniques	Reconstruction consistency; Visual inspection
5	synth. PET/MR	State-of-the-art techniques	Visual inspection; SUV line profiles

5.3.1 Compared Methods

In the following the compared methods for this chapter will be discussed.

Manifold Alignment-Based

The examined methods included a two-step method as described above, two variations of a one-step method, and several variations of the groupwise method. The compared manifold alignment methods were:

- SGA.SIM: The simultaneous groupwise manifold alignment (SGA) approach using the image similarity-based kernel proposed in Eq. (5.5).
- SGA.REG: The SGA approach using the registration-based similarity kernel proposed in Eq. (5.9).
- SGA.NN: The SGA approach using the image similarity based kernel from Eq. (5.5), with a conventional nearest neighbour sparsification instead of the global bipartite maximum edge similarity matching proposed above.
- SGA.LEM: The SGA approach using a Laplacian Eigenmaps-based intra-dataset kernel. The extension of LEM to two datasets is very similar to the LLE-extension. It can be obtained by replacing the LLE cost function (i.e. $\phi(\cdot)$) in Eq. (4.6) by the LEM cost function (see Belkin and Niyogi [2003]) and following through the derivations in Appendix A. The inter-dataset kernel was the same as for SGA.SIM.
- ICP: The basic two-step method as described in Section 5.2.2.
- FULL: The basic one-step approach as described in Section 5.2.2. For this method all similarity matrices U_{pq} between all slice positions \mathcal{C}_p and \mathcal{C}_q were defined using the image-based similarity kernel defined in Eq. (5.5).
- JMR: Joint manifold representation is a one-step manifold alignment approach proposed by Torki et al. [2010], and discussed earlier in Section 3.5. The intra-dataset kernels are defined using LEM, and the intra-dataset kernels using an orthogonalised SSD similarity measure.

Methods Not Based on Manifold Alignment

In addition to the manifold alignment-based techniques two state-of-the art slice stacking techniques from the literature were investigated. Note that both methods were discussed in more detail in Section 3.1.4.

- IMBASED: Dikaios et al. [2012] proposed a slice-stacking technique which finds corresponding 2D MR slices across slice positions by directly evaluating a similarity measure (mutual information).
- PBNV: Würslin et al. [2013] used a simultaneously acquired MR pencil beam navigator to establish respiratory correspondences between 2D MR slice-by-slice data. Thus each acquired slice had an associated navigator value. To reconstruct volumes, in the original work, the navigator values were used to bin the MR data from different slice positions into four gates. For the experiments here this approach was adapted

to allow for more accurate reconstructions. In the reconstruction step, for each slice a volume was reconstructed by selecting the data from all other slice positions with navigator values closest to the input slice.

Note that the slice stacking technique proposed by Von Siebenthal et al. [2007a] is not included in the evaluations. That method requires an acquisition protocol in which a fixed navigator slice is acquired before and after each data slice, and could therefore not be implemented using the data used here.

5.3.2 Comparison of Manifold Alignment Methods

Experiment 1: Validation on Synthetic Data

For the validation on synthetic data, synthetic 2D MR slice-by-slice images were generated from real scans of 10 healthy volunteers as described in Section B.1 of the Appendix. As for the real slice-by-slice acquisitions used later on, for each volunteer slices from 29-34 overlapping slice positions with 50 images per slice position were obtained.

In order to evaluate the accuracy of the various reconstruction techniques, a leave-one-out (LOO) cross validation was performed for each subject separately. For one of the synthetic slices the whole volume it belonged to was left out, apart from the slice itself. Next, a new volume was reconstructed around the slice from the remaining synthetic data, which resulted in an approximation \hat{V}_{out} of the left-out ground-truth volume V_{out} . The reconstruction error was estimated by calculating the \mathcal{L}_2 -distance $\mathcal{L}_2(\hat{V}_{out}, V_{out})$, between the two volumes. The process was repeated for each of the synthetic slices to obtain a good estimate of the mean and the standard deviation of the reconstruction error.

In preliminary evaluations SGA.SIM and SGA.REG were found to be relatively insensitive to the choice of the parameters k (number of neighbours in the LLE cost function) and σ (the kernel shape parameter) as long as they were chosen from a reasonable range. The number of neighbours was set to 70% of the time points acquired at each slice position, i.e. $k = 35$, and $\sigma = 0.5$ for the evaluation of both methods. The significance of the weighting parameter μ which governs the importance of the similarity kernel in Eq. (4.6) and the embedded dimensionality d of the manifold representation were more closely investigated.

In order to estimate the inherent dimensionality of the slice-by-slice data a technique developed for this purpose which employs fractal dimensions to estimate the underlying number of dimensions of high-dimensional data was employed [Camastra and Vinciarelli, 2002]. The algorithm was applied to the data from each slice position and each subject and an average was computed. The average inherent dimensionality of the data was found to be close to 3 and hence the embedding dimensionality was set to $d = 3$.

The optimal parameter value of μ for each of the methods using this parameter⁴ and the results of the LOO cross validation were determined simultaneously in a 2-fold nested cross validation over all subjects. That is, the 10 subjects were divided into a tuning set of 5 and a test set of 5. The LOO experiment was performed on the tuning set for a range of different parameter values of μ , combining the errors of the 5 subjects into a single error figure for each parameter. The parameter range was chosen to be 0.001 to 4544 with a logarithmic spacing⁵. Next, the μ associated with the minimum error for each method was selected and used to evaluate the reconstruction error for each of the subjects in the test set. Lastly, the positions of the test set and tuning set were swapped to evaluate the second fold.

Experiment 2: Validation on Real Data

In order to validate the proposed method on real data, slice-by-slice data was acquired from 10 volunteers over the course of 24 to 28 minutes as described in Section 5.2.1. The resulting data typically consisted of 29-32 slice positions, each sampled 50 times. Note that all volunteers except volunteer H were the same as the volunteers from Experiment 1. Volunteer H was not available for the long scanning session and was replaced by volunteer K. Additionally, a 1D pencil beam navigator was recorded for each slice. However, only a leading pencil beam navigator was acquired since acquiring both leading and trailing navigators was not possible on the scanner for an acquisition of this length. For each of the examined manifold alignment methods a volume was reconstructed around each of the acquired slices resulting in a time sequence of 3D volumes.

Since for the real data the ground-truth is unknown, the consistency of reconstruction was evaluated for each of the techniques. That is, for each acquired slice a volume V was reconstructed, and for each slice s in this volume another volume \hat{V}_s was reconstructed. Ideally, \hat{V}_s should be equal to V . In practice, however, a different input slice position will yield a different reconstruction. To estimate this reconstruction consistency the \mathcal{L}_2 -distance $\mathcal{L}_2(\hat{V}_s, V)$ was calculated for each slice of each reconstructed volume. Consistency alone naturally does not give an indicator of the correctness of a method, as reconstructions can be consistently wrong. However, together with visual inspection of the results, and the results from the experiment on synthetic data from the previous section, the consistency error gives an indication as to how reproducible the shown results are.

Because for the real data no ground-truth was available, there was no good measure to use for tuning the values of the parameters. Applying the dimensionality estimation technique to the real data in the same manner as before yielded an approximate average

⁴That is, FULL, JMR, SGA.NN, SGA.LEM, SGA.SIM, and SGA.REG.

⁵This unusual upper bound number is due to the fact that the original range had to be extended in logarithmic steps until it could be confirmed that the error of SGA.LEM converged.

underlying dimensionality of 5 for each slice position. However, in preliminary experiments it was found that none of the manifold alignment techniques were robust to matching in such high dimensions. This may have to do with the fact that higher modes of variation were slice position specific and not common to all slice positions as with the first three modes. Therefore, the same dimensionality as for the synthetic experiment was used, that is $d = 3$. The respective μ -parameters for each method were set to the values found in the synthetic tuning step. The other parameters were set to the same values as for the synthetic experiments, i.e. $\sigma = 0.5$, $k = 25$.

5.3.3 Comparison to State-of-the-Art Slice-Stacking Methods

Based on the findings of Experiments 1 & 2, the two best performing methods, i.e. SGA.SIM and SGA.REG, were compared to the two slice stacking techniques not based on manifold alignment, i.e. IMBASED and PBNVAV.

Experiment 3: Validation on Synthetic Data

The experimental set up was exactly the same as for Experiment 1, except for the compared methods.

Experiment 4: Validation on Real Data

Also for the real data no changes were made in the experimental protocol on real data from Experiment 2, except the examined methods were replaced.

Experiment 5: PET/MR Simulation

In the final experiment, IMBASED, SGA.SIM and SGA.REG were applied for the retrospective respiratory motion correction of a realistic synthetic PET dataset using the reconstruct-transform-average (RTA) approach, which was discussed in Section 3.1.4.

Synthetic PET/MR data was generated from real 4D CT and real MR data from one of the healthy volunteers as discussed in Section B.2 of the Appendix. 4D CT data was used in order to obtain very realistic motion fields inside of the lungs. The resulting synthetic PET data consisted of reconstructions from 30 gates with relatively low counts in each gate. Furthermore, a synthetic lesion was placed in one of the lungs close to the diaphragm. An example of a single gate is shown in Figure B.1d in the Appendix. The synthetic MR slice-by-slice data for this experiment consisted of 29 slice positions, each of which was sampled 30 times.

In order to perform motion correction of the PET data using the slice-by-slice MR data, a volume was reconstructed from a synthetic slice for each time point using the IMBASED,

SGA.SIM, and SGA.REG methods. Again, the parameters for the latter two were chosen based on the findings in Experiment 1. The dimensionality was set to $d = 3$ and the weight parameter was set to $\mu = 0.144$ for SGA.SIM and SGA.REG. Furthermore, $\sigma = 0.5$ and $k = 21$, which was again 70% the time points acquired at each slice position. Note that performing this experiment using PBNV, by extracting an artificial pencil beam navigator signal from the images, would not be meaningful. This is because all of the slices originating from one volume would have identical navigator values and the matching would always be perfect.

Similar to the acquisition of the real slice-by-slice data (see Section 5.2.1), the input slice position was varied for each time point by the square root of the total number of slice positions. For each of these reconstructed volumes backward motion fields were derived by registering the reconstructed volumes, obtained using the three examined techniques, to the MR reference volume using B-spline registration [Modat et al., 2010].

Finally, each of the synthetic PET gates was transformed back to the reference exhale state using the estimated backwards motion fields derived using each of the examined techniques. The transformed PET gates were then averaged to produce the final PET image. Additionally, the PET gates without motion corruption were averaged to obtain a ground-truth, and a PET volume without any motion correction was reconstructed. In order to compare the quality of motion correction, intensity profiles through the tumour were evaluated for each of the reconstructions above.

5.4 Results

5.4.1 Comparison of Manifold Alignment Methods

Experiment 1: Validation on Synthetic Data

The combined error scores on the synthetic tuning sets for each parameter value of μ and each of the manifold alignment methods are shown in Figure 5.5. It can be observed that the groupwise techniques, SGA.SIM and SGA.REG, outperform the other techniques for almost all parameter values μ . Furthermore, it can be seen that suitable parameter values for both of these techniques can be chosen from a range of values between approximately 0.05 and 0.5.

The results of the leave-one-out cross-validation on the respective synthetic data test sets are shown in Table 5.2.

Since the error distribution of all volunteers was symmetric but not normal, the significance levels were evaluated using a 1-tailed Wilcoxon signed rank test. It was observed that overall SGA.REG performed significantly better ($p < 0.001$) than all other evaluated techniques, closely followed by SGA.SIM. SGA.LEM also performed well compared to the

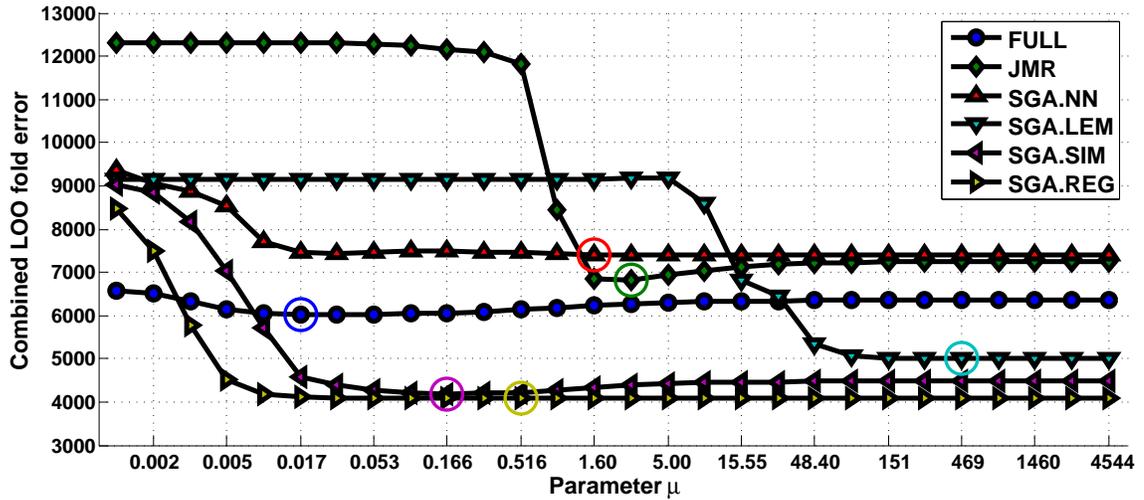


Figure 5.5: Combined fold error on the tuning set for the examined manifold alignment methods which depend on the parameter μ . The respective minimum of each method is indicated by a larger coloured circle.

other methods, but did not reach the reconstruction accuracy of SGA.SIM and SGA.REG. When examining results volunteer-by-volunteer, it was found that for each volunteer the improvements of SGA.REG over all the other techniques were significant ($p < 0.001$). The average reconstruction error over all volunteers on synthetic data using each of the examined methods is shown in Figure 5.6.

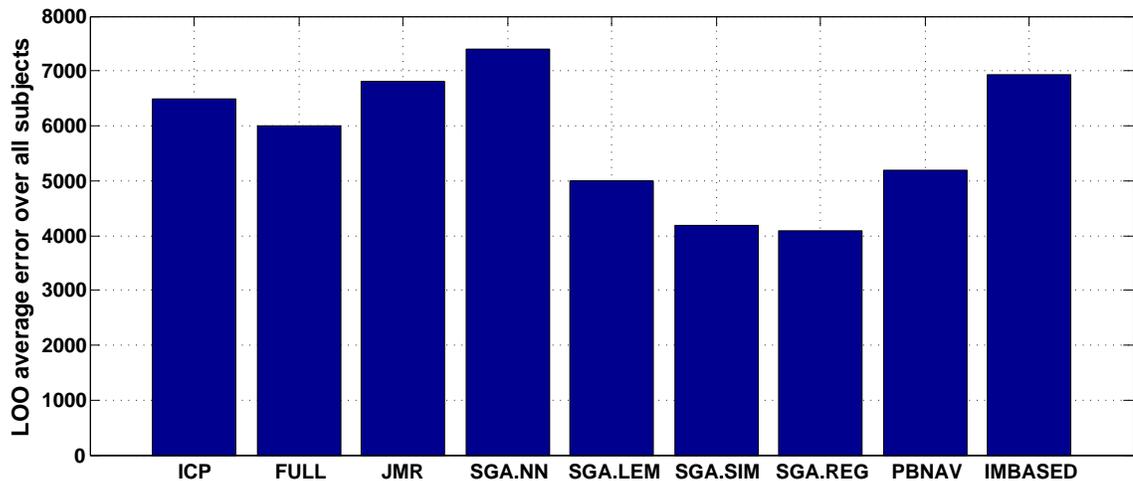


Figure 5.6: Average error over all volunteers for all methods evaluated in the two experiments on synthetic data (Exp. 1 and 3).

In order to visualise the reconstructed volumes a maximum intensity projection (MIP) in the R-L direction over the left lung was used. Examples of such MIPs of volumes reconstructed for volunteer E using the examined manifold alignment techniques, and their

disparity with respect to the ground-truth, are shown in Figure 5.7. The top row (EXH) contains a typical end-exhale reconstruction and the bottom row (INH) a typical inhale.

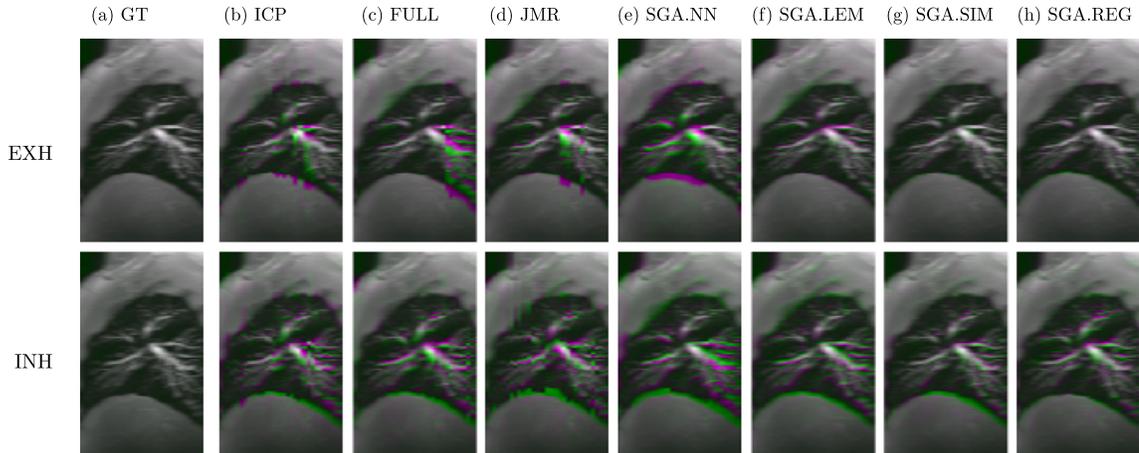


Figure 5.7: Examples of sagittal MIPs over the left lung of volunteer E obtained from synthetic slice-by-slice data reconstructed using the examined techniques. The left most column (a), contains the ground-truth (GT) MIP. The remaining columns show the disparity between the GT and the volumes reconstructed using the examined techniques: (b) ICP, (c) FULL, (d) JMR, (e) SGA.NN, (f) SGA.LEM, (g) SGA.SIM, (h) SGA.REG. Areas where the GT and the reconstruction have the same value are shown in grey. Areas where the intensities deviate from the GT are shown in magenta or green. The projections are shown at two respiratory positions: end-exhale (EXH), and end-inhale (INH).

Visual inspection of the results confirmed the results in Table 5.2. The reconstructions using SGA.REG most closely match the ground-truth. While SGA.SIM and SGA.LEM also gave reasonable reconstructions they don't match the accuracy of SGA.REG.

Experiment 2: Validation on Real Data

In Figure 5.9 sagittal MIPs of volumes reconstructed from real slice-by-slice data using the manifold alignment-based reconstruction techniques are presented for one representative volunteer (volunteer B). Reconstructions are shown at three distinct respiratory states (end-exhale, mid-inhale, and end-inhale).

By visually inspecting the reconstructions on real data it can be seen that most of the techniques performed well in reconstructing end-exhale respiratory states. This was presumably due to the fact that end-exhale states are more reproducible than other respiratory states and that most volunteers spend more time at this respiratory position, which led to more data being acquired there. For other respiratory states SGA.REG yielded the most realistic reconstructions out of the examined manifold alignment methods. While the reconstructions obtained using SGA.SIM also look reasonable, SGA.LEM performed

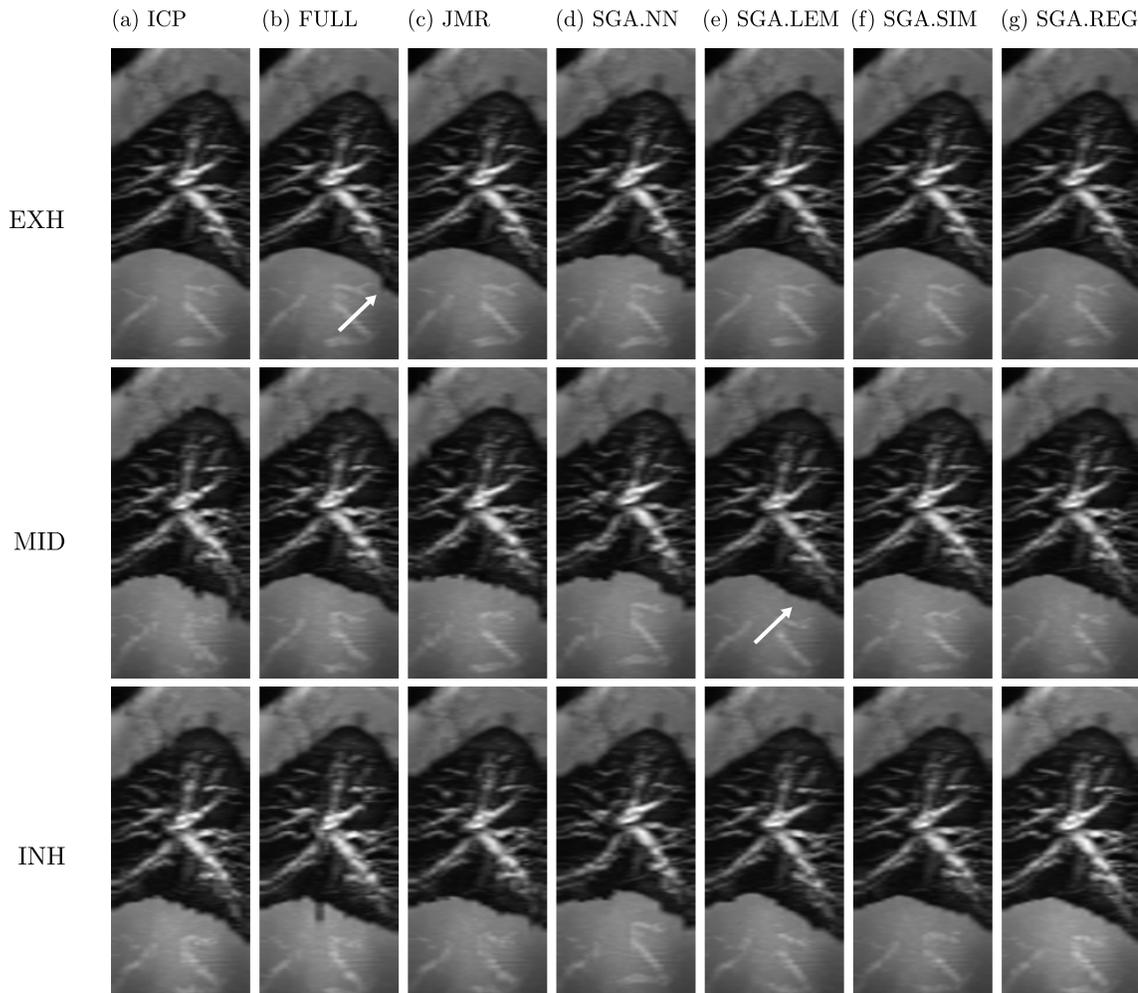


Figure 5.8: Examples of sagittal MIPs over the left lung of volunteer B obtained from real slice-by-slice data reconstructed using the examined manifold alignment methods: (a) ICP, (b) FULL, (c) JMR, (d) SGA.NN, (e) SGA.LEM, (f) SGA.SIM, (g) SGA.REG. The projections are shown at three respiratory positions: end-exhale (EXH), mid-inhale (MID), at end-inhale (INH). White arrows indicate some of the reconstruction errors.

significantly poorer on the real data.

The results of the consistency experiment are shown in Table 5.3. For the majority of volunteers SGA.LEM produced the most consistent results, closely followed by SGA.REG. Both the improvements of SGA.LEM over SGA.REG and the improvements of SGA.REG over the rest of the methods were statistically significant ($p < 0.001$), except for volunteer K, for whom SGA.LEM, SGA.REG and FULL performed similarly. Again, statistical significance was assessed using a 1-tailed Wilcoxon signed rank test as the error distributions were symmetric but not normal.

5.4.2 Comparison to State-of-the-Art Slice-Stacking Methods

After establishing that robust 4D MR reconstructions can be obtained using manifold alignment techniques, the two best performing methods, SGA.SIM and SGA.REG, were evaluated against two slice stacking techniques which do not rely on manifold learning. Note that the numerical results of SGA.SIM and SGA.REG are the same here as for Experiments 1 and 2 and are only replicated here for comparison.

Experiment 3: Validation on Synthetic Data

The results of the leave-one-out cross-validation on the synthetic data test sets for the comparison against the state-of-the-art methods are shown in Table 5.4.

SGA.REG performed significantly better ($p < 0.001$) than all other evaluated techniques, and SGA.SIM performed significantly better than PBNV and IMBASED. When examining results volunteer-by-volunteer it was found that for 7 out of the 10 subjects the improvements of SGA.REG over the respective next best technique (highlighted in light grey in Table 5.4) were significant ($p < 0.001$). For volunteers C, G and I, SGA.REG was significantly better than PBNV and IMBASED. The average performance over all volunteers of PBNV and IMBASED is also included in the comparison in Figure 5.6.

The error figures observed were also reflected by the visual inspection of MIPs through the reconstructed volumes. In Figure 5.9, such MIPs are shown for all techniques at end-inhale and end-exhale respiratory states.

All examined techniques were able to reconstruct accurate volumes at end-exhale. However, SGA.REG consistently yielded more accurate reconstructions at other respiratory states when compared to the other techniques. The reconstructions using PBNV suffered from artefacts in the anterior chest-lung interface (indicated by arrows in the top row of Figure 5.9). The biggest improvements of SGA.REG over the other techniques were achieved for inhale states, such as the one shown in the bottom row of Figure 5.9, where breathing variabilities are largest. The purely image-based technique, IMBASED, occasionally contained large misestimations of the respiratory position such as, for example, shown for the end-inhale state (INH) in Figure 5.9.

Experiment 4: Validation on Real Data

Figure 5.10 shows sagittal MIPs of volumes reconstructed from real slice-by-slice data using the two state-of-the-art reconstruction techniques and the groupwise manifold alignment methods. Examples of the same volunteer already presented in Experiment 2 (volunteer B) are shown at three distinct respiratory states (end-exhale, mid-inhale, and end-inhale). For the mid-inhale and end-inhale reconstructions a small area inside the lungs was magnified

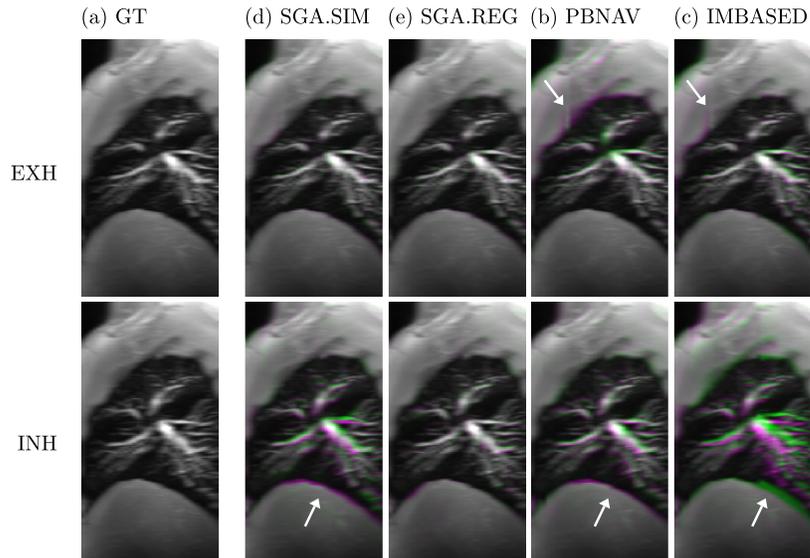


Figure 5.9: Examples of sagittal MIPs over the left lung of volunteer E obtained from synthetic slice-by-slice data reconstructed using the four examined techniques. The left most column (a), contains the GT MIP. The remaining columns show the disparity between the ground-truth (shown in magenta) and the volumes reconstructed using the examined techniques: (b) PBNV, (c) IMBASED, (d) SGA.SIM, (e) SGA.REG (shown in green). The projections are shown at two respiratory positions: (EXH) end-exhale, and (INH) deep inhale. The white arrows indicate some of the disparities.

to highlight the differences in vessel continuity.

As in the previous experiment visual inspection of these results revealed that end-exhale respiratory states were reconstructed accurately by all techniques. Only IMBASED sometimes suffered from implausible reconstructions such as the kink in the diaphragm indicated by an arrow in the EXH row in Fig. 5.10. For other respiratory positions the effects of misalignment became more evident. In mid-inhale frames, such as the one shown in the MID row of Figure 5.10, occasionally inaccuracies in the anterior chest-lung interface were observed for PBNV (highlighted by an arrow) similar to the ones observed on synthetic data. The vessel close-up views show that SGA.REG gave the smoothest estimate in this area of all examined techniques. Artefacts like the ones mentioned above became much worse at inhale states, such as the one shown in the bottom row of Figure 5.10 where respiratory variabilities are the largest. In this frame PBNV suffered from large artefacts on the diaphragm and the anterior chest-lung interface as highlighted by arrows, which were likely due to respiratory intra- and inter-cycle variations. Also the vessels were distorted and intermittent for PBNV while SGA.REG retained smooth continuous vessels. As can be seen in the same figure, IMBASED and SGA.SIM sometimes gave anatomically implausible reconstructions at inhale.

The results of the consistency experiment are shown in Table 5.5. For all volunteers PBNV gave significantly ($p < 0.001$) more consistent results than all other evaluated techniques, and SGA.REG gave significantly more consistent results than IMBASED and SGA.SIM. Again, significance was assessed using a 1-tailed Wilcoxon signed rank test as the error distributions were symmetric but not normal. The fact that PBNV achieved higher consistency is not surprising and will be further discussed in Section 5.5.

To illustrate the process of manifold alignment, the first two dimensions of the aligned embedding using SGA.REG of $\mathcal{G}^{(18)}$ (group 18), which contains the manifolds of slice position 18 (blue), and 19 (red), are shown in Figure 5.11b. For comparison also the embeddings of the same two slice positions computed directly from LLE without any alignment are shown in Figure 5.11a. The embedded low-dimensional coordinates which correspond to the slices used for slice positions 18 and 19 of the inhale (INH) reconstruction using SGA.REG in Figure 5.10 are highlighted with black squares in Figures 5.11a and b. In Figure 5.10 the two slice positions used to generate the embeddings in Figure 5.11 are indicated by small red arrows in the bottom row.

Experiment 5: PET/MR Simulation

The results of the PET motion correction experiment are shown in Figures 5.12 and 5.13. A coronal slice including the tumour through the ground-truth reconstruction without motion corruption is shown in Figure 5.12a and the reconstruction without motion correction is shown in Figure 5.12b. The motion corrected reconstructions using IMBASED, SGA.SIM, and SGA.REG are shown in Figures 5.12c, 5.12d, and 5.12e, respectively. Close-up views of the tumour for (a)-(e) are shown in Figure 5.12f. Finally, the SUV line profiles through the tumour along the line indicated in Figure 5.12a, are shown in Figure 5.13.

From Figure 5.12 it can be seen that the reconstructions using SGA.SIM (Figure 5.12d) and SGA.REG (Figure 5.12e) both very closely matched the ground-truth with no motion corruption (Figure 5.12a). The reconstruction using IMBASED (Figure 5.12c) was markedly less defined. This is confirmed by looking at the line profiles through the tumour shown in Figure 5.13. The profile through the SGA.REG reconstruction was very similar to the motionless ground-truth. The profile from the SGA.SIM reconstruction was slightly less defined.

5.5 Discussion

The purpose of this chapter was to investigate the utility of manifold alignment for the purpose of navigator-less, accurate reconstruction of high-resolution, high-contrast dynamic 3D volumes from long slice-by-slice 2D MR acquisitions. In particular a novel approach based

on the simultaneous groupwise manifold alignment of adjacent slice positions (SGA.REG) has been proposed which outperforms state-of-the-art slice stacking techniques proposed for PET/MR motion correction.

The evaluation has aimed to answer two different research questions. First, a number of manifold alignment approaches were compared in terms of reconstruction accuracy and consistency in Experiments 1 and 2. It was found that the groupwise techniques (in particular SGA.SIM and SGA.REG) significantly outperformed the two-step manifold alignment approach (ICP), and the conventional one-step approaches (FULL and JMR). Naive two-step approaches, such as the ICP approach, do not take anatomical similarities into account and a matching purely on the shape of the embedded manifolds lacks robustness. Furthermore, the shape of the embeddings may vary by more than an affine transformation in some cases. In addition to the issues regarding the robustness of the matching, two-step approaches require determining the correct flipping and ordering of the eigenvectors prior to alignment, which may not always be feasible. For those reasons, for the remainder of this thesis the two-step approach was not further explored. The one-step approaches FULL and JMR required calculating similarities between all slice positions. Not surprisingly, this led to very inaccurate matches for anatomically distant slice positions explaining the poor performance of those methods. The groupwise approaches overcome this limitation by only comparing adjacent slice positions, for which some anatomical similarity can be expected. This is particularly true for overlapping slice positions such as those used in this chapter.

Additionally, different versions of the groupwise approach were explored. The SGA.SIM approach is considered to be the base algorithm for the scope of this discussion. Different attributes were removed from it or added to investigate their effect. The benefits of the novel sparsification technique based on a bipartite node matching (see Section 5.2.3) were explored through a comparison to a normal nearest neighbour-based sparsification (SGA.NN). Both on real and on synthetic data the proposed sparsification technique proved significantly superior to SGA.NN, and the sparsification proved to be one of the most important components for accurate reconstructions. Furthermore, a groupwise approach extending LEM instead of LLE was explored. The resulting method, SGA.LEM, proved slightly inferior to SGA.SIM in the synthetic experiments. It is interesting to note that SGA.LEM performed at its best at much higher parameter values of μ than all other methods (see Fig. 5.5), which also explains its good performance in the consistency experiment. This means that a lot of trust is placed on the inter-dataset similarity kernel, rather than the inherent shape of the manifold. Even though the reconstructions are consistent it can easily be verified in the visual inspection of the real reconstructions in Fig. 5.8 that the reconstructions are of poor quality. Based on these findings, for the remainder of this thesis LLE based approaches will be used. Lastly, a more accurate similarity kernel accounting for remaining anatomi-

cal differences between adjacent slice positions (SGA.REG) was investigated. The method produced significantly more accurate reconstructions than the other methods for both real and synthetic data. It is expected that the improvements would be even more pronounced if non-overlapping or thicker slices would be acquired. However, SGA.REG required the computation of the expensive kernel in Eq. (5.9), which took up to 3 hours per subject running in parallel on 8 cores on a workstation clocked at 2.7 GHz. In contrast, SGA.SIM ran in a few seconds. Thus, the choice between SGA.SIM and SGA.REG depends on the requirements of the application.

The second part of the evaluations (Experiments 3 and 4) focused on a comparison of SGA.REG and SGA.SIM to PBNV and IMBASED, which do not employ manifold alignment. On the synthetic data both SGA.SIM and SGA.REG significantly outperformed PBNV and IMBASED. Visual inspection of the real reconstructions in Experiment 4 confirmed these findings. Manifold alignment techniques capture the breathing motion at each slice position in a low-dimensional manifold embedding. For a sufficiently high dimensionality of the embeddings, this representation of motion has the potential to model all respiratory inter-cycle as well as intra-cycle variabilities such as amplitude variations, baseline shifts and hysteresis in the manifold representation. Such variabilities are then taken into account during the reconstruction resulting in more accurate volumes than current methods. PBNV relies on a 1-dimensional signal on the diaphragm and thus cannot capture such variabilities. Although the reconstructions were accurate in the diaphragm region, artefacts were present in other parts of the anatomy such as the lung-chest interface highlighted by white arrows in Figures 5.9 and 5.10. This is due to the fact that although the same point may pass through a navigator window on the diaphragm, the overall organ configuration may vary due to inter- and intra-cycle variabilities [Blackall et al., 2006].

IMBASED should in principle be able to capture respiratory variabilities as it relies only on comparisons of the high-dimensional image data. As for the groupwise manifold alignment approaches only adjacent slice positions were directly compared. The performance of the IMBASED technique showed that relying only on high-dimensional similarities is not sufficient for accurate slice stacking. SGA.SIM and SGA.REG only used such similarities to guide the alignment of the inherent manifold structure of each slice position, which led to much better results.

The consistency results of Experiment 4 shown in Table 5.5 allow quantification of how consistently the visual results shown in Figure 5.10 can be reproduced over all time frames. The consistency results should however not be mistaken as a measure for reconstruction accuracy and should not be interpreted on their own. PBNV gave the most consistent results for all volunteers. This can be understood when considering that a 1D signal is easier to match consistently to 1D signals from other slice positions. This, however, comes

at the cost of not being able to model inter- and intra-cycle breathing variabilities, which cannot be captured by a 1D signal only. The other techniques require matchings in higher dimensions which have more ambiguity, but allow for such variabilities to be captured. The high consistency of PBNV also means that errors due to mismatches caused by breathing variabilities were also consistently reproduced. SGA.REG exhibited a high consistency for all volunteers when compared to SGA.SIM and IMBASED and offered a good compromise between reconstruction reproducibility and ability to model complex breathing patterns.

Acquiring data using a slice-by-slice acquisition protocol as described in Section 5.2.1 allowed excellent image contrast and made it possible to image vessel structures inside the lungs which cannot be visualised using a dynamic 3D MR acquisition protocol. As has been demonstrated such motion fields can be used to accurately correct for motion in simultaneous PET/MR acquisitions. Capturing respiratory variabilities is not yet an issue due to the comparatively low image resolutions currently achieved in PET. However, PET has the potential to acquire images with an image resolution down to 0.67 mm [Moses, 2011]. In this case, very accurate motion estimation will gain importance.

Coronal slices were used because in this manner fewer slice positions are needed to cover the whole thorax and because anatomical changes are smaller from slice position to slice position. In Section 2.1 it was discussed that most of the respiratory motion occurs in the S-I and A-P directions, therefore using sagittal input slices may be beneficial. However, the manifold alignment methods discussed in this chapter are not robust to sagittal input slices because of the rapidly changing anatomy. However, note that through-plane motion is also captured in the manifold which allows stacking slices with the same through-plane motion. In the PET/MR scenario motion is not estimated from individual 2D MR slices but rather from reconstructed 3D volumes in which through-plane motion should be accounted for. A remaining problem is that as a consequence of the coronal input slices the reconstructed volumes have a lower resolution of 4 mm in the A-P direction compared to 1.4 mm in the R-L and S-I directions, which may impede motion estimation accuracy. Therefore, in the following two chapters strategies are explored to extend the methodology in this chapter to sagittal input slices.

For the PET reconstructions the RTA approach was used. It has been shown by Polycarpou et al. [2011] that under some circumstances the MCIR method may provide superior PET motion correction than RTA, but this comes at greater computational cost. The technique proposed in this chapter can also be used to provide motion estimates for MCIR-based reconstructions.

During a PET/MR session typically many different MR scans need to be acquired for, amongst other things, attenuation correction and visualisation of different aspects of the anatomy. Note that the MR-based PET motion correction scheme proposed here could

be applied with minimal scanning overheads, since acquiring one 2D slice for 160 *ms* per cardiac cycle is sufficient to retrospectively obtain high-contrast 3D volumes for the entire duration of a PET imaging session. In the remainder of the time the scanner could be used for other MR imaging relevant to the PET session. In the future it would also be conceivable to integrate part of those additional scanning requirements into the proposed method, for example, by temporally interleaving partial MR data acquisitions using different protocols and reconstructing them retrospectively using SGA.REG.

4D MR volumes, such as the ones reconstructed in this chapter, have many potential applications in addition to PET/MR motion correction. For example, they may be useful for the planning of treatments such as radiotherapy or image-guided treatments, they may be used as input for motion models, or as a tool to better understand and quantify respiratory motion.

Table 5.2: Means and standard deviations of the leave-one-out cross-validation of the manifold alignment based methods on synthetic slice-by-slice data for each volunteer. The method giving the lowest reconstruction error is highlighted in dark grey for each volunteer; the next best method is highlighted in light grey.

	ICP $\cdot 10^3$	FULL $\cdot 10^3$	JMR $\cdot 10^3$	SGA.NN $\cdot 10^3$	SGA.LEM $\cdot 10^3$	SGA.SIM $\cdot 10^3$	SGA.REG $\cdot 10^3$
Volunteer A	3.932 \pm 1.328	3.125 \pm 0.979	3.909 \pm 1.324	4.747 \pm 2.089	3.459 \pm 1.500	2.627 \pm 0.926	2.508 \pm 0.902
Volunteer B	4.811 \pm 1.760	4.918 \pm 1.342	5.786 \pm 1.556	8.652 \pm 2.207	4.700 \pm 1.653	3.818 \pm 0.955	3.774 \pm 0.970
Volunteer C	4.693 \pm 1.962	4.980 \pm 1.679	5.863 \pm 1.961	5.786 \pm 2.502	4.522 \pm 1.814	3.685 \pm 1.180	3.649 \pm 1.161
Volunteer D	5.526 \pm 2.541	5.949 \pm 1.996	6.183 \pm 2.041	7.582 \pm 2.247	3.905 \pm 2.055	3.478 \pm 1.909	2.988 \pm 0.959
Volunteer E	6.434 \pm 2.287	8.812 \pm 2.500	9.426 \pm 2.077	8.825 \pm 2.824	5.738 \pm 1.711	4.830 \pm 1.142	4.673 \pm 1.200
Volunteer F	7.867 \pm 1.663	10.043 \pm 3.130	11.190 \pm 2.983	9.747 \pm 2.246	7.675 \pm 3.055	6.192 \pm 1.717	6.094 \pm 1.642
Volunteer G	6.554 \pm 2.64	5.843 \pm 2.165	6.741 \pm 2.269	6.406 \pm 1.988	5.930 \pm 2.460	5.189 \pm 2.308	5.105 \pm 2.247
Volunteer H	4.935 \pm 1.819	6.122 \pm 2.118	6.494 \pm 2.201	8.606 \pm 2.582	4.535 \pm 1.859	3.557 \pm 1.042	3.465 \pm 1.009
Volunteer I	4.798 \pm 1.298	5.026 \pm 1.353	6.362 \pm 1.457	7.580 \pm 1.487	4.937 \pm 1.641	3.859 \pm 1.269	3.806 \pm 1.271
Volunteer J	4.217 \pm 2.128	5.250 \pm 2.276	6.216 \pm 2.189	6.012 \pm 2.204	4.594 \pm 1.839	3.372 \pm 0.816	3.372 \pm 0.816

Table 5.3: Means and standard deviations of the consistency experiment using the manifold alignment methods on real slice-by-slice data for each volunteer. The method giving the lowest consistency error for each volunteer is highlighted in dark grey, and the method with the second to lowest consistency error is highlighted in light grey.

	ICP $\cdot 10^3$	FULL $\cdot 10^3$	JMR $\cdot 10^3$	SGA.NN $\cdot 10^3$	SGA.LEM $\cdot 10^3$	SGA.SIM $\cdot 10^3$	SGA.REG $\cdot 10^3$
Volunteer A	3.467 \pm 1.331	3.039 \pm 1.461	3.501 \pm 1.529	5.219 \pm 3.293	1.184 \pm 2.074	3.238 \pm 2.473	2.706 \pm 2.080
Volunteer B	4.529 \pm 1.894	4.278 \pm 2.115	4.470 \pm 1.788	7.227 \pm 4.558	2.893 \pm 3.057	3.975 \pm 2.910	3.416 \pm 2.281
Volunteer C	5.152 \pm 1.883	4.393 \pm 1.899	5.347 \pm 1.904	7.715 \pm 4.902	2.931 \pm 3.670	4.616 \pm 3.112	3.969 \pm 2.658
Volunteer D	5.026 \pm 2.387	4.875 \pm 2.419	5.257 \pm 2.280	7.479 \pm 4.572	3.845 \pm 4.183	4.687 \pm 3.552	4.258 \pm 3.137
Volunteer E	6.218 \pm 2.492	5.191 \pm 2.411	6.342 \pm 2.586	7.683 \pm 5.286	4.655 \pm 4.601	4.650 \pm 3.096	4.295 \pm 2.860
Volunteer F	7.631 \pm 3.017	7.280 \pm 3.653	7.489 \pm 2.986	8.136 \pm 5.489	5.228 \pm 5.515	6.026 \pm 3.865	5.963 \pm 3.824
Volunteer G	6.826 \pm 2.856	5.865 \pm 2.454	7.126 \pm 3.000	9.127 \pm 5.059	5.331 \pm 5.467	6.495 \pm 4.514	5.638 \pm 3.649
Volunteer I	5.518 \pm 2.207	4.405 \pm 2.102	5.340 \pm 2.082	7.080 \pm 4.413	3.672 \pm 3.848	4.537 \pm 2.851	4.090 \pm 2.699
Volunteer J	4.312 \pm 1.817	4.134 \pm 1.778	5.052 \pm 2.041	6.630 \pm 4.298	3.940 \pm 3.648	4.189 \pm 2.549	3.723 \pm 2.407
Volunteer K	6.416 \pm 2.422	5.022 \pm 2.109	6.715 \pm 2.691	8.027 \pm 5.678	4.866 \pm 4.629	5.075 \pm 3.279	4.832 \pm 3.061

Table 5.4: Means and standard deviations of the leave-one-out cross-validation of manifold alignment methods compared to state-of-the-art methods on synthetic slice-by-slice data for each volunteer. The method giving the lowest reconstruction error is highlighted in dark grey for each volunteer; the next best method is highlighted in light grey.

	SGA.SIM .10 ³	SGA.REG .10 ³	PBNAV .10 ³	IMBASED .10 ³
Volunteer A	2.627 ± 0.926	2.508 ± 0.902	2.911 ± 1.215	4.604 ± 2.150
Volunteer B	3.818 ± 0.955	3.774 ± 0.970	3.940 ± 1.113	7.802 ± 2.993
Volunteer C	3.685 ± 1.180	3.649 ± 1.161	3.907 ± 1.472	4.822 ± 1.543
Volunteer D	3.478 ± 1.909	2.988 ± 0.959	3.183 ± 1.031	6.146 ± 2.966
Volunteer E	4.830 ± 1.142	4.673 ± 1.200	5.441 ± 2.129	7.995 ± 2.773
Volunteer F	6.192 ± 1.717	6.094 ± 1.642	7.076 ± 2.650	9.043 ± 3.002
Volunteer G	5.189 ± 2.308	5.105 ± 2.247	5.861 ± 2.705	5.941 ± 2.041
Volunteer H	3.557 ± 1.042	3.465 ± 1.009	4.184 ± 1.698	7.653 ± 3.173
Volunteer I	3.859 ± 1.269	3.806 ± 1.271	4.044 ± 1.357	6.723 ± 2.282
Volunteer J	3.460 ± 0.851	3.372 ± 0.816	3.967 ± 1.294	5.820 ± 2.232

Table 5.5: Means and standard deviations of the consistency experiment on real slice-by-slice data for each volunteer. The method giving the lowest consistency error for each volunteer is highlighted in dark grey, and the method with the second to lowest consistency error is highlighted in light grey.

	SGA.SIM .10 ³	SGA.REG .10 ³	PBNAV .10 ³	IMBASED .10 ³
Volunteer A	3.238 ± 2.473	2.706 ± 2.080	1.777 ± 1.632	5.730 ± 3.671
Volunteer B	3.975 ± 2.910	3.416 ± 2.281	2.181 ± 1.842	6.960 ± 4.278
Volunteer C	4.616 ± 3.112	3.969 ± 2.658	2.850 ± 2.160	7.807 ± 4.327
Volunteer D	4.687 ± 3.552	4.258 ± 3.137	2.566 ± 2.513	8.717 ± 5.700
Volunteer E	4.650 ± 3.096	4.295 ± 2.860	2.585 ± 2.321	8.123 ± 5.137
Volunteer F	6.026 ± 3.865	5.963 ± 3.824	4.026 ± 2.959	8.560 ± 5.140
Volunteer G	6.495 ± 4.514	5.638 ± 3.649	4.166 ± 2.942	8.598 ± 4.447
Volunteer I	4.537 ± 2.851	4.090 ± 2.699	2.807 ± 1.995	8.309 ± 5.256
Volunteer J	4.189 ± 2.549	3.723 ± 2.407	2.766 ± 2.069	6.712 ± 3.933
Volunteer K	5.075 ± 3.279	4.832 ± 3.061	3.841 ± 2.600	8.636 ± 5.528

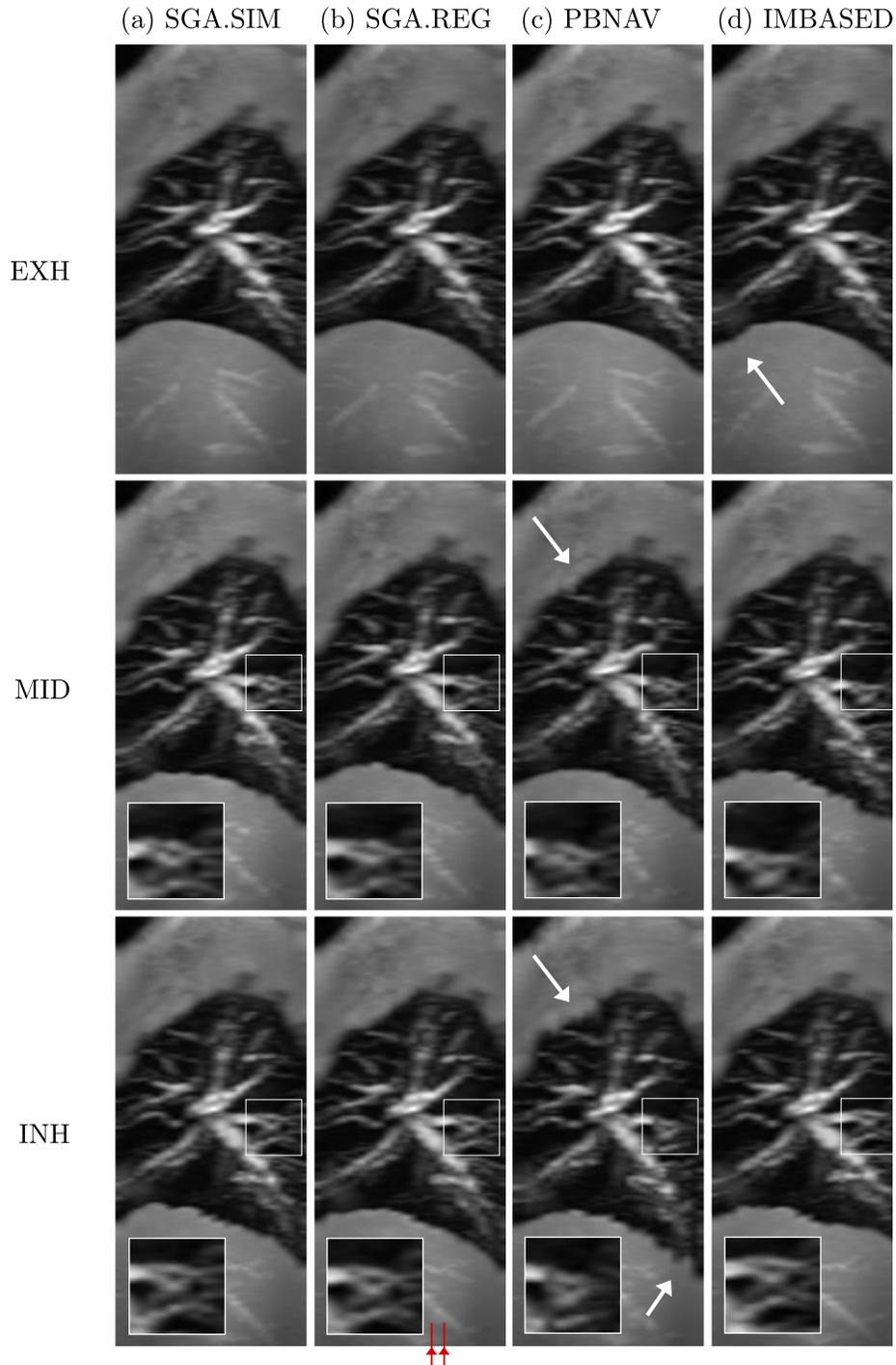


Figure 5.10: Examples of sagittal MIPs over the left lung of volunteer B obtained from real slice-by-slice data reconstructed using the four examined methods: (a) SGA.SIM, (b) SGA.REG, (c) PBNVAV, and (d) IMBASED. The projections are shown at three respiratory positions: end-exhale (EXH), mid-inhale (MID), at end-inhale (INH). Magnifications of the square area indicated with white boxes in the middle and bottom rows are shown in the lower left corner of the respective images. The red markers in the lower right mark the slice positions shown in the sample manifold embedding in Figure 5.11.

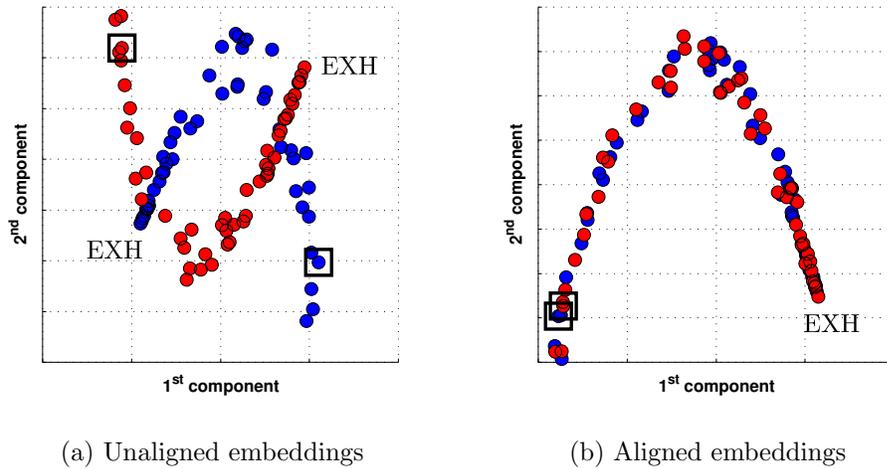


Figure 5.11: Unaligned and aligned embeddings for slice positions 18 (blue) and 19 (red) of volunteer B. (a) shows the first two dimensions of the 3-dimensional embeddings of the two slice positions as obtained directly from LLE without alignment. (b) shows the first two dimensions of group 18 of the aligned SGA.REG embedding. The manifold coordinates of the slices that were chosen for the inhale reconstruction of SGA.REG in Figure 5.10 are highlighted with black rectangles.

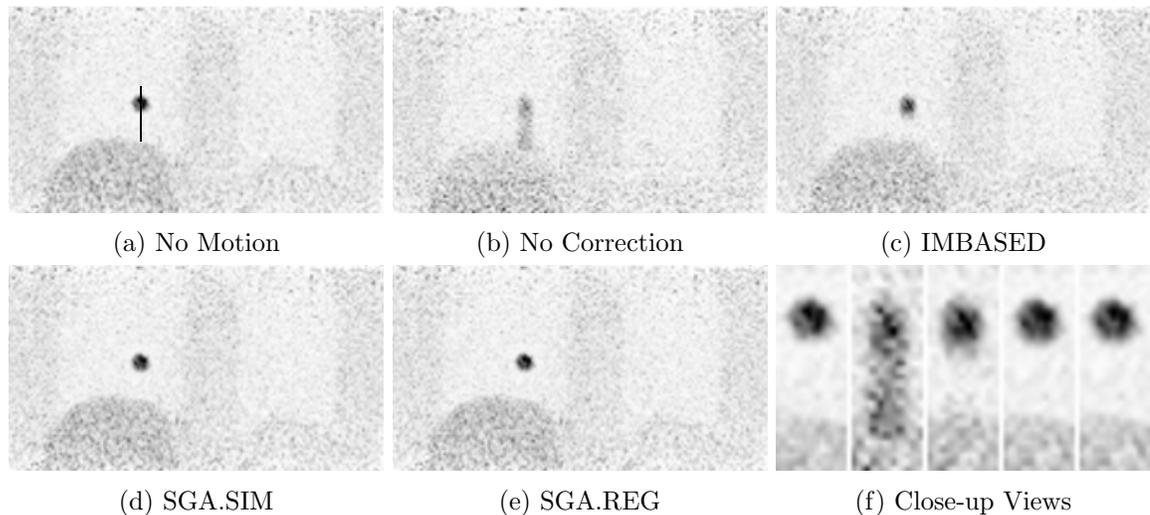


Figure 5.12: Result of motion correction applied to the synthetic PET data: (a) The sum of the PET gates with no motion (ground-truth), (b) PET reconstruction without motion correction, and motion corrected PET reconstructions using (c) IMBASED, (d) SGA.SIM and SGA.REG. Finally close-up views of the tumour for (a)-(e) are shown in (f). The line profiles in Figure 5.13 were calculated along the dotted line in (a).

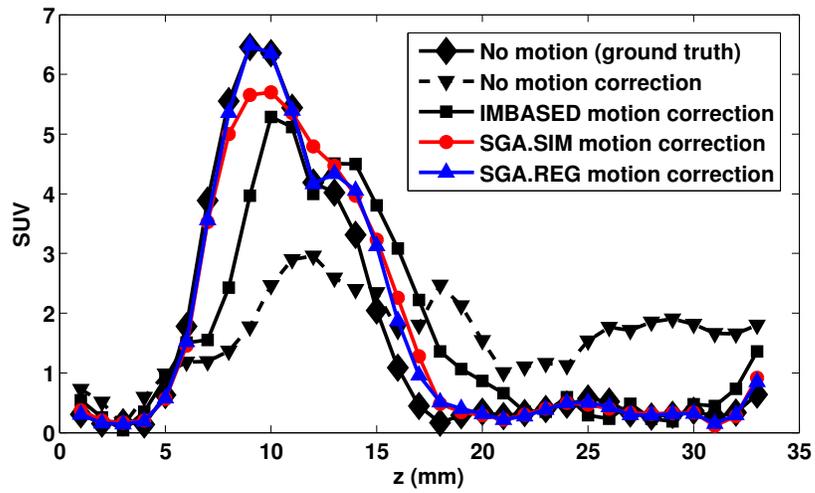


Figure 5.13: SUV line profiles through the tumour along the line indicated in Figure 5.12a for the motionless ground-truth, no motion correction, and motion correction using IM-BASED, SGA.SIM and SGA.REG. The motionless ground-truth and the SGA.REG motion correction have almost identical profiles between $z = 5 \text{ mm}$ and $z = 12 \text{ mm}$.

Chapter 6

Autoadaptive Motion Modelling

In this chapter manifold alignment will be investigated to build a motion model which can automatically adapt to new breathing patterns while capturing respiratory inter- and intra-cycle variabilities.

6.1 Introduction

In interventions such as HIFU which require, or are aided by, image-based planning prior to the treatment, motion induced by respiration may cause the target area to deviate from the treatment plan, reducing the accuracy of the treatment and potentially causing harm to the patient [Keall et al., 2006; Tanter et al., 2007; Tzifa et al., 2010].

Recent advances in MR compatible materials and equipment now allow interventions such as percutaneous cardiac interventions [Tzifa et al., 2010] and HIFU [Hynynen et al., 1996] to be performed inside an MR scanner. MR-guidance has also been proposed for radiotherapy [Stam et al., 2012; Raaymakers et al., 2009], but so far it has never been performed on humans. In such interventions, in addition to the aforementioned problem of target motion, essential intra-procedural imaging data are affected by respiratory motion. An example of such images are temperature maps derived from MR phase images, which are used for guiding HIFU treatments. Respiratory motion results in phase variations and may lead to temperature artefacts [Köhler et al., 2011]. Both the problem of target motion and the problem of imaging artefacts create the need for tools that can robustly estimate 3D motion from MR data. Similarly, hybrid systems such as PET-MR scanners [Zaidi and Del Guerra, 2011] and hybrid MR linear accelerators [Stam et al., 2012] can greatly benefit from motion estimations that can be derived from MR data.

6.1.1 Motion Modelling

One solution to the problem of estimating respiratory motion is the use of a motion model, in which the patient-specific breathing motion is learned before the treatment and used to correct for respiratory motion during the treatment. As was discussed in detail in Section 3.3, motion models typically consist of three distinct stages:

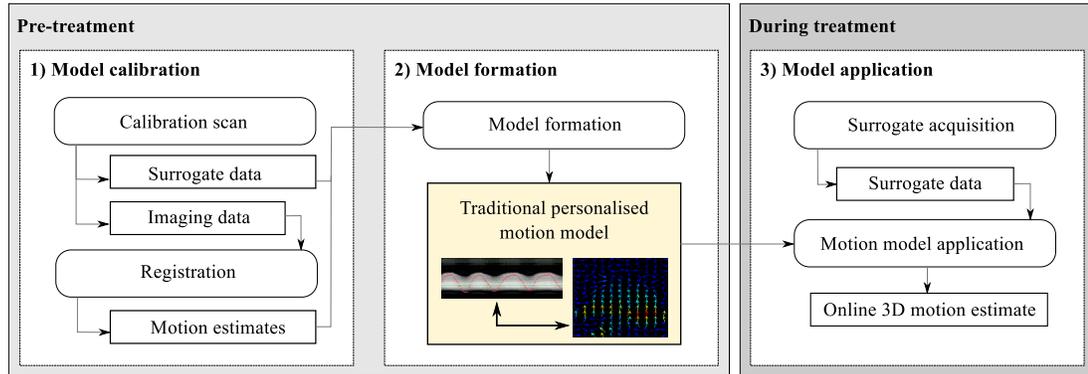
1. Model Calibration: Acquisition of imaging and surrogate data and estimation of the motion.
2. Model Formation: relating the motion estimations to the surrogate data.
3. Model Application: motion estimation during the treatment from the surrogate data alone.

This traditional motion model paradigm is also outlined in Figure 6.1a.

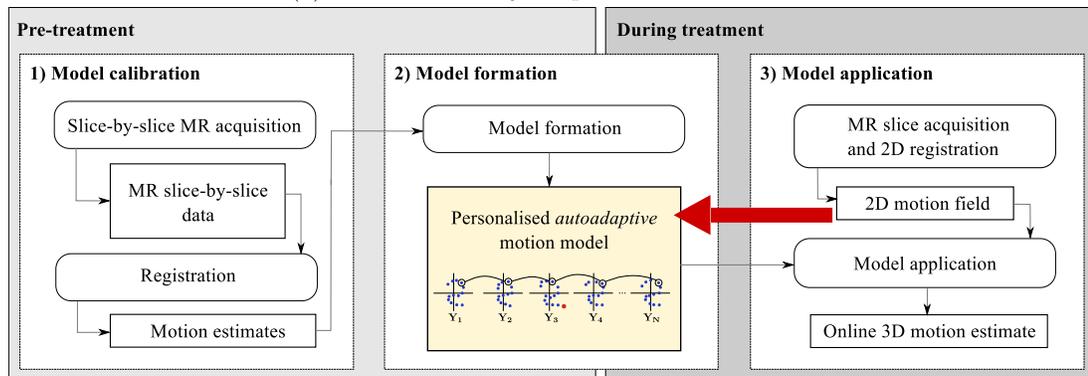
An underlying assumption of the majority of traditional motion models is that the nature of the relationship between the surrogate data and the motion (i.e. the correspondence model) remains constant. However, for long treatment durations it is possible for the breathing motion to undergo significant changes, for example due to varying degrees of relaxation of the patient during the procedure, because of pain or discomfort experienced [King et al., 2012; Hoogeman et al., 2009] or because of organ drift [Arnold et al., 2011]. In the traditional motion model paradigm the model is formed before the treatment and has no ability to adapt to changing breathing patterns.

In response to this problem, a small number of papers have proposed *adaptive* motion modelling techniques, which were discussed in Section 3.3.3. These systems are always characterised by the availability of the calibration imaging data also during the treatment, e.g. tumour locations in fluoroscopy-guided radiotherapy [Schweikard et al., 2000], or MR data in PET/MR acquisitions [King et al., 2012]. Most of these methods require rebuilding of the motion model periodically based on new imaging data, i.e. by repeating steps 1) and 2) in Fig. 6.1a.

To the best of my knowledge, the only motion modelling technique which has the potential to automatically adapt to new breathing patterns without the need to recalibrate the model was proposed by Isaksson et al. [2005] for radiotherapy. However, the objective of that work was to track a single point of interest (i.e. the tumour). In this chapter, an *autoadaptive* motion modelling technique is proposed which can be used to obtain full motion estimates in the thorax and abdomen. This is achieved by altering the traditional motion model paradigm as shown in Fig. 6.1b. In the proposed framework the data acquired during the model calibration and application stages are of the same type, but are only required to contain partial motion information. In particular, motion fields estimated



(a) Traditional subject-specific motion model



(b) The proposed autoadaptive motion model

Figure 6.1: Schematic representation of (a) the traditional subject-specific motion paradigm and (b) the autoadaptive subject-specific motion model allowing for continuous adaptivity to changing breathing patterns. The red arrow indicates the proposed change to the motion model paradigm allowing for new surrogate/calibration data to be incorporated into the motion model without interrupting the application phase. In this new paradigm the model is initially formed pre-treatment, but is updated continually during the treatment.

from 2D MR slices are used both to form the 3D motion model, and as surrogate data to apply it. This allows the surrogate data acquired in the application phase to be fed back into the model formation phase as the treatment goes on, allowing a continuous updating of the model by partial motion information. In Fig. 6.1b this update process is indicated by the red feedback arrow.

In the following it is demonstrated how such a motion model can be implemented using manifold alignment. In particular, the SGA technique, which was proposed in the previous Chapter, is extended to build an autoadaptive motion model from multiple 2D motion fields derived from 2D MR slices acquired at different anatomical positions. These 2D motion fields can then be combined to estimate full 3D motion in the thorax. To allow this, significant changes to the original SGA methodology were necessary:

- Manifold alignment was applied to dense 2D motion fields rather than the image intensities.
- Because through-plane motion can cause significant registration artefacts in coronal slices, predominately sagittal 2D MR slices were chosen as the calibration data. However, SGA is not robust to only sagittal input slices. Therefore, the method was extended to use slices of both coronal and sagittal orientations.
- In order to estimate motion at points where the respiratory pattern was not yet sufficiently sampled, a new interpolation strategy was developed.

The remainder of this Chapter is structured as follows. In Section 6.2 it will be shown how SGA can be extended to be robustly applied as an autoadaptive motion model. In Section 6.3 the technique will be evaluated on synthetic and real data. Finally, Section 6.4 contains the discussion.

This work was presented in Baumgartner et al. [2014b] in a preliminary form. A journal submission covering most of the material in this Chapter is currently under review.

6.2 Materials and Methods

SGA was originally proposed for coronal input slices, since the anatomy changes less from slice position to slice position in this plane. In the motion modelling context, however, it is essential that the input data captures as much of the motion as possible. It is well known that respiratory motion is largest in the S-I and A-P directions (see Chapter 2). In order to capture the majority of respiratory motion and avoid through-plane motion artefacts, in this Chapter sagittal input slices were used. Unfortunately, SGA as described in the previous chapter is not robust to sagittal input slices, because respiratory information often gets lost while propagating from group to group through the body centre, where anatomy changes rapidly from slice position to slice position and there is less respiratory motion. Therefore, here the technique is extended to additionally incorporate data acquired from a single coronal slice position to aid this transition through the body centre.

In the following the proposed autoadaptive motion modelling approach is outlined by following the three motion modelling stages. First it is shown how sagittal and coronal input slices are acquired and 2D motion estimates are derived. Next, the extension of SGA to use motion fields, rather than images is discussed, and it is shown how different slice orientations are incorporated into the model. Lastly, it is shown how the model can be updated during a treatment in the application phase, and how this leads to continuous adaptivity.

6.2.1 Calibration Scan

Similar to the acquisition scheme in Chapter 5, the entire region of interest (i.e. the entire thorax) is divided into adjacent slice positions $\mathcal{S}_1, \dots, \mathcal{S}_L$ each spanning 8 mm. The two differences with respect to the protocol described in Section 5.2.1 are that (1) the slice orientation is sagittal, and (2) the slice positions are not overlapping in order to speed up the acquisition. Additionally, data is acquired from one coronal slice position \mathcal{C} to help the proposed algorithm with the propagation of respiratory information between distant sagittal slices. A schematic of the slice positions is shown in Fig. 6.2. The coronal slice is chosen such that it coincides with the dome of the left hemi-diaphragm in order to maximise the amount of captured respiratory motion.

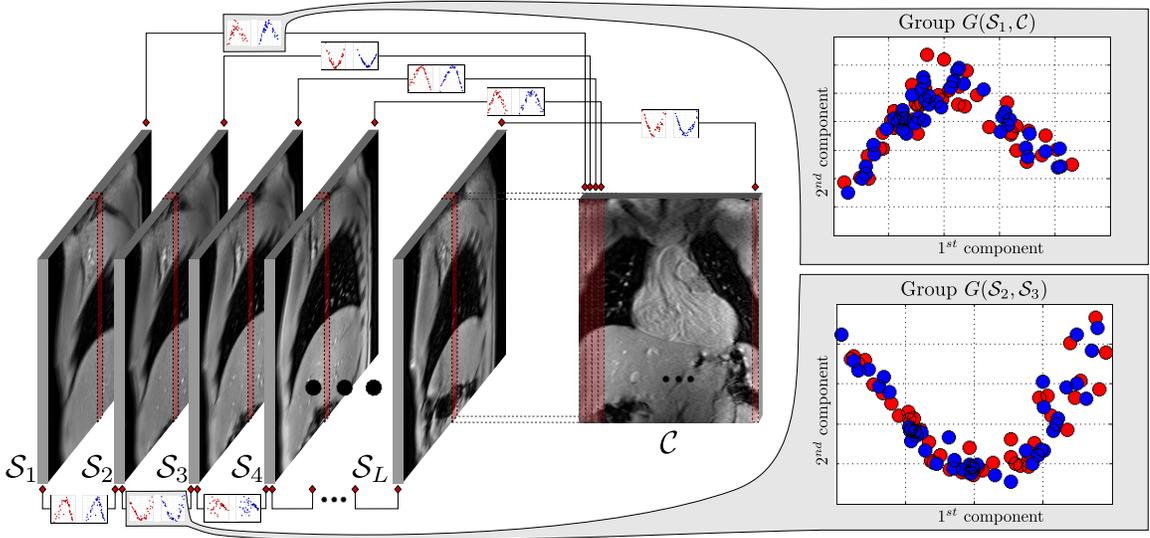


Figure 6.2: Schematic of group connections through simultaneous manifold embeddings. Motion fields from neighbouring and orthogonal slices can be embedded simultaneously using appropriate similarity kernels, leading to aligned embeddings. Two close-up views of aligned manifold embeddings, originating from a dataset with 50 motion fields per slice position, are shown on the right.

2D images b_p^i are acquired from these slice positions in a slice-by-slice fashion, iterating through the slice positions, first the sagittals then the coronal, until each slice position is covered τ_p times. As before, only one slice is acquired per heart beat at systole in order to isolate the respiratory motion and maximise the contrast of the input slices. The acquisitions were carried out on a Philips Achieva 3T MR scanner using a T1-weighted gradient echo sequence with an acquired in-plane image resolution of $1.4 \times 1.4 \text{ mm}^2$ and a slice thickness of 8 mm. To cover the entire thorax typically around 30 sagittal slice positions were needed. Additionally, exhale slices $b_p^{(exh)}$ were acquired using the same slice-by-slice protocol in a scan consisting of two consecutive breath-holds. The volunteers were

instructed to try and reproduce the same exhale position as best as they could. Lastly, a 1D pencil beam navigator was acquired immediately before each dynamic image solely for the purpose of validating the method.

In the next step, 2D motion fields c_p^i are derived for each slice position by registering each of the τ_p 2D images b_p^i to the corresponding slice $b_p^{(exh)}$ from the exhale breath-hold image. The NiftyReg implementation [Modat et al., 2010] of a non-rigid B-spline registration algorithm [Rueckert et al., 1999] with 3 hierarchy levels, a final grid spacing of 15 mm in each direction and no bending energy penalty term is used. The vectorised motion fields c_p^i derived from the slice positions \mathcal{S}_p and \mathcal{C} , respectively, form the datasets $X_{(p,sag)}$ and $X_{(cor)}$.

6.2.2 Motion Model Formation

A groupwise embedding of all the motion data acquired during the calibration phase can be viewed as a surrogate-driven motion model as it contains all respiratory information collected during the calibration and can be applied using new 2D motion information as the surrogate data, as will be explained in Section 6.2.3. Thus, in order to form the motion model an embedding of all the sagittal and coronal slices acquired during the calibration phase is performed in groups of two as described in Sections 4.3 and 5.2.2.

In order to embed the motion data derived from sagittal and coronal motion fields, two significant methodological novelties need to be introduced. First, new similarity kernels of the form described in Eq. (4.8) are needed with which motion fields of slices with the same as well as slices with different orientations can be compared. In particular, appropriate distance functions $\epsilon(\cdot, \cdot)$ need to be defined for both of these cases. Secondly, a new propagation scheme is needed which allows respiratory information to propagate across the body centre.

Distance Functions for Neighbouring Slices of the Same Orientation

The choice for neighbouring sagittal motion data is related to the registration based similarity kernel proposed in the previous chapter in Section 5.2.3 for coronal images. Here it is adapted to motion fields. For two neighbouring slice positions \mathcal{S}_p and \mathcal{S}_q the distance of data points $x_{(p,sag)}^i$ and $x_{(q,sag)}^j$ is assessed based on the \mathcal{L}_2 -distance of the corresponding motion fields $c_{(p,sag)}^i$ and $c_{(q,sag)}^j$. In order to account for the changes in anatomy between sagittal slices, one of the motion fields is transformed into the coordinate system of the other using transformations $T_{q \rightarrow p}, T_{p \rightarrow q}$ which are obtained by registering the breath-hold slices $b_{(p,sag)}^{(exh)}$ to $b_{(q,sag)}^{(exh)}$, and vice versa, and then correcting them for non-translational ef-

fects using the method proposed in Rao et al. [2002]¹. As before in Chapter 5, to increase robustness, the results of the comparisons in the spaces of slice positions \mathcal{S}_p and \mathcal{S}_q are averaged. The final distance measure is defined as

$$\begin{aligned} \epsilon_{neighb.}(x_{(p,sag)}^i, x_{(q,sag)}^j) = & \frac{1}{2} \mathcal{L}_2(c_{(q,sag)}^i, T_{p \rightarrow q}(c_{(p,sag)}^j)) + \\ & \frac{1}{2} \mathcal{L}_2(c_{(p,sag)}^i, T_{q \rightarrow p}(c_{(q,sag)}^j)). \end{aligned} \quad (6.1)$$

In Chapter 5 it was shown that including the transformations $T_{q \rightarrow p}, T_{p \rightarrow q}$ significantly improves the matching accuracy compared to the simple \mathcal{L}_2 -distance between images. Since the changes from sagittal slice position to sagittal slice position can be even larger than for coronal slices it is expected that this effect will be more pronounced for slices of this orientation.

Distance Function for Slices with Different Orientation

To define a distance function for two slices acquired from a sagittal slice position \mathcal{S}_p and a coronal slice position \mathcal{C} , the fact is used that such slices have an overlap and thus visualise the same anatomy in the overlapping region as is illustrated in the example in Fig. 6.3a. Motion estimates derived from two such slices share the S-I motion component along the slice overlap. If $q_{(p,sag)}^i$ is the S-I motion in the overlapping region originating from the i -th acquired sagittal slice at \mathcal{S}_p and $q_{(cor)}^j$ the motion originating from the j -th acquired coronal slice, the similarity function is defined as

$$\epsilon_{orthog.}(x_{(p,sag)}^i, x_{(cor)}^j) = \mathcal{L}_2(q_{(p,sag)}^i, q_{(cor)}^j). \quad (6.2)$$

To illustrate this, examples of S-I line motions $q_{(p,sag)}^i$ and $q_{(cor)}^j$ originating from sagittal and coronal slices are shown in Fig. 6.3b. The left hand side shows S-I motion extracted along the intersection (highlighted in Fig. 6.3a) from a coronal slice and the curves on the right hand side show two example S-I motions extracted from the same region from the sagittal slice position. The blue curve shows a good match in respiratory position of the sagittal to the coronal slice and will lead to a low distance in Eq. (6.2). Conversely, the motion in the sagittal slice from which the red curve was extracted has a higher distance to the coronal slice motion and thus corresponds to a different respiratory state.

¹Note that the registrations between slice positions of the breath-hold volumes only have to be performed once for the model formation, but not during the model application.

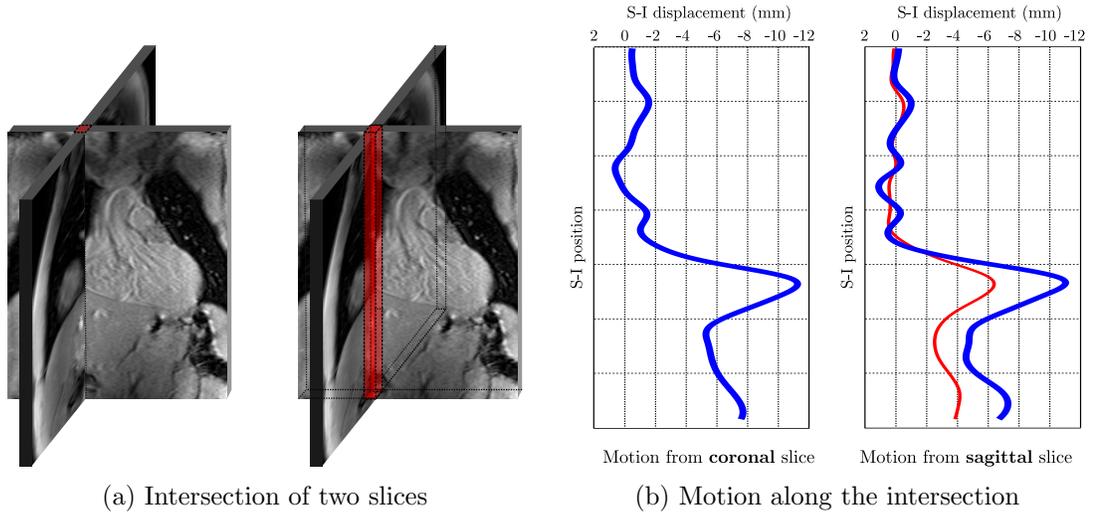


Figure 6.3: Derivation of similarity kernel based on motion in slice overlap. (a) Illustration of the overlap of two orthogonal slices, (b) S-I motion components derived from the overlapping area from a coronal (left) and sagittal (right) slice position. Two possible S-I motion components originating from the sagittal slice are shown: One that is similar to the one derived from the coronal slice (blue) and hence corresponds to a similar motion state, and one that is dissimilar (red) and consequently corresponds to a different motion state.

Group Connectivity and Propagation of Respiratory Information

By using the distance measures defined in Eqs. (6.1) and (6.2) it is now possible to simultaneously embed any two neighbouring sagittal slice positions \mathcal{S}_p and \mathcal{S}_q and any overlapping sagittal and coronal slice positions \mathcal{S}_p and \mathcal{C} . This is achieved by converting the distances into similarities using Eq. (4.8) and then solving the general optimisation in Eq. (4.9) to obtain an embedding.

This allows a manifold embedding to be performed using data from all acquired slice positions in overlapping groups of two. The groups are formed as follows: all neighbouring sagittal slice positions embedded in overlapping groups, and additionally the data from each sagittal slice position is embedded together with the data from the coronal slice position. This is illustrated in Fig. 6.2. By embedding the data in this way, the 2D motion fields from all slice positions are embedded in three groups, with the exception of \mathcal{S}_1 and \mathcal{S}_L which don't have a left-hand or right-hand neighbour, respectively. For example, data from slice position \mathcal{S}_3 is embedded in the groups $\mathcal{G}(\mathcal{S}_2, \mathcal{S}_3)$, $\mathcal{G}(\mathcal{S}_3, \mathcal{S}_4)$ and $\mathcal{G}(\mathcal{S}_3, \mathcal{C})$.

Note that the data within each group are aligned, as is illustrated by the close-up views of $\mathcal{G}(\mathcal{S}_1, \mathcal{C})$ and $\mathcal{G}(\mathcal{S}_2, \mathcal{S}_3)$ in Fig. 6.2.

As opposed to the SGA approach from the previous chapter, there now no longer is just one path from each slice position to each other slice position. Rather, the different slice positions are now connected by a network of groups as is illustrated in Fig. 6.4.

This is a crucial element of the proposed technique which allows propagation of respiratory information in the form of low-dimensional embedded coordinates from slice position to slice position without having to go through difficult areas such as the body centre where there are larger anatomical differences between adjacent slices.

In the following section it will be shown how low-dimensional coordinates obtained from a 2D input motion field can be propagated to all slice positions.

6.2.3 Model Updating and Adaptivity

After the calibration scan and model formation phase the model is ready to be applied. During the application phase slices can be acquired in the same slice-by-slice fashion as described in Section 6.2.1. That means that each input image is acquired at a different slice position and can have sagittal or coronal orientation.

From each of these slices a new 2D motion estimate can be derived and embedded in the groups containing data from this slice position. This then allows reconstruction of a pseudo 3D motion estimate by looking up corresponding 2D motions from all other slice positions. Note that the resulting 3D motion fields will lack the R-L motion component. The new 2D motion, as well as being used as the surrogate input to the motion model, is retained in the manifold embeddings of the appropriate groups. This leads to the desired autoadaptivity. Each of these steps will be explained in detail below. The process is illustrated using the example shown in Fig. 6.4.

Obtaining a 2D Update Motion Estimate

In a first step, the most recently acquired image $b_p^{(new)}$ is registered in 2D to the corresponding breath-hold exhale slice $b_p^{(exh)}$ in order to form the current updated motion field $c_p^{(new)}$. That is, the motion field $c_p^{(new)}$ obtained in this way acts as the surrogate data for the motion model application. The same registration parameters as in the initial calibration are used (see Section 6.2.1). On a workstation with 8 cores clocked at 2.7 GHz this operation took around 500 *ms*. In the example in Fig. 6.4, it is assumed that the newest slice is acquired at slice position \mathcal{S}_2 , which is highlighted in yellow. Note that only the registrations from $b_p^{(new)}$ to $b_p^{(exh)}$ have to be performed during the application phase. The registrations across slice positions (i.e. $b_p^{(exh)}$ to $b_q^{(exh)}$), which are required for the registration based similarity kernel described in Section 6.2.2, only need to be performed once during the model formation.

Obtaining Corresponding 2D Motion Fields for All Slice Positions

In order to estimate 3D motion from partial motion information provided by the single input 2D motion field $c_p^{(new)}$, the motion from the newly acquired slice must be related

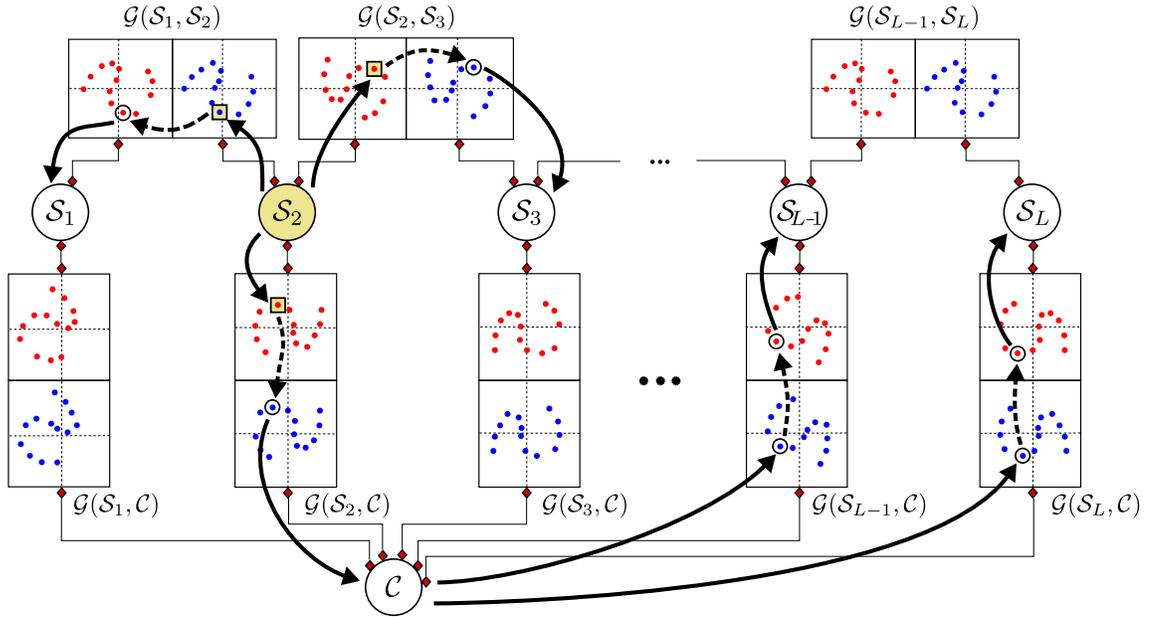


Figure 6.4: Schematic of the connection of slice positions by means of pairwise embedding and propagation of respiratory information through the manifolds. Assuming a new input slice at \mathcal{S}_2 , the neighbouring groups can be directly updated as indicated by the squares with yellow background. Then, through a combination of nearest neighbour searches (dotted arrows) and group transitions based on shared data (solid arrows), low-dimensional coordinates which correspond to the respiratory state can be propagated to all remaining slice positions.

to that from all other slice positions. First, all groups which contain data from the slice position at which the new slice is acquired must be recalculated. If, as in the example in Fig. 6.4, the current update slice is acquired at slice position \mathcal{S}_2 , the groups $\mathcal{G}(\mathcal{S}_1, \mathcal{S}_2)$, $\mathcal{G}(\mathcal{S}_2, \mathcal{S}_3)$ and $\mathcal{G}(\mathcal{S}_2, \mathcal{C})$ must be re-evaluated. To achieve this the dataset X_2 is simply augmented by the new entry and the respective embeddings are recalculated. Updating just a few groups is very fast and on average took less than 100 *ms* in a single threaded MATLAB implementation.

As is shown in Fig. 6.4, the new motion field now has a corresponding low-dimensional point in each of the low-dimensional embeddings which include dataset X_2 . These points are highlighted by squares with yellow backgrounds in Fig. 6.4. The coordinates of these low-dimensional embedded points are propagated from group to group following the shortest path, i.e. using the path requiring the fewest group transitions. This is done by making use of the fact that the groups share datasets, and that the datasets within a group are aligned. This effectively means that neighbouring slices are updated through the sagittal-to-sagittal groups and further away sagittal slice positions are connected through the coronal slice. Other methods for choosing the update paths taking into account the quality of

the embedding were also investigated and may lead to small improvements. However, the improvements gained may not justify the additional complexity.

Following the example in Figure 6.4, first the nearest neighbours of low-dimensional points corresponding to data from \mathcal{S}_1 , \mathcal{S}_3 and \mathcal{C} are found in the respective groups. The nearest neighbour operation is indicated by the dotted arrows in Fig. 6.4 and the nearest points are indicated by circles. The high-dimensional motion fields corresponding to the circled points in $\mathcal{G}(\mathcal{S}_1, \mathcal{S}_2)$ and $\mathcal{G}(\mathcal{S}_2, \mathcal{S}_3)$ are at the same respiratory position as the input slice. Note that in the example in Fig. 6.4, only one nearest neighbour is shown per slice position. In reality κ nearest neighbours are identified at this stage and the corresponding 2D motion fields are interpolated. This is described in more detail in the following section.

Next, motion fields from all other sagittal slice positions are chosen by using the coronal slice. The nearest low-dimensional neighbour of the input point in group $\mathcal{G}(\mathcal{S}_2, \mathcal{C})$ can then be transported to all other groups containing the coronal slice because that same point exists in all other groups. Note that only the closest neighbour is transported across groups. From there the corresponding points from the sagittal motion datasets are again found by looking up the κ nearest neighbours.

At the end of this process κ 2D motion fields have been identified for each slice position. In the following section, it will be described how these 2D motion fields can be combined to arrive at an interpolated motion estimate for each slice position, and how these partial 2D motion estimates can then be stacked into a full pseudo 3D motion field.

Interpolating Motion Fields on the Manifold and 3D Reconstruction

If the motion model has not yet fully sampled all the possible motion states of the new breathing pattern, it is important that it has the ability to interpolate between the motion states which are already there.

In order to estimate the 2D motion field for a slice position, κ nearest neighbours are identified for each slice position as described in the previous section. The estimated motion field is then given as a weighted average of the κ motion fields corresponding to those nearest neighbours. That is, the estimated motion field for a slice position q is given by

$$c_q^{(est)} = \frac{\sum_{i \in \eta(y_p^j)} s_i c_q^i}{\sum_{i \in \eta(y_p^j)} s_i}, \quad (6.3)$$

where $\eta(y_p^j)$ are the κ nearest neighbours on the manifold of slice position q to the low-dimensional point y_p^j from a slice position p which is sharing a group with q . Furthermore, $s_i = \frac{1}{\omega_i}$, where ω_i is distance of each neighbour to y_p^j in the manifold embedding. This process is illustrated in the close-up of $\mathcal{G}(\mathcal{S}_2, \mathcal{S}_3)$ which is shown in Figure 6.5.

The $c_p^{(est)}$ for all sagittal slice positions $\mathcal{S}_1, \dots, \mathcal{S}_L$ are then stacked into a pseudo 3D

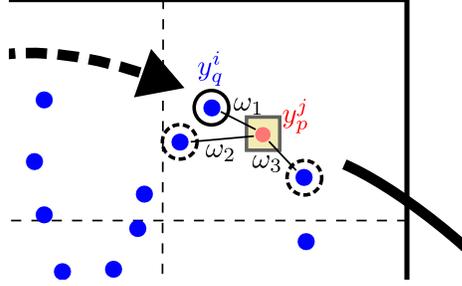


Figure 6.5: Close-up view of the right-hand side of group $\mathcal{G}(\mathcal{S}_2, \mathcal{S}_3)$ in Figure 6.4. Here $p = \mathcal{S}_2$ and $q = \mathcal{S}_3$. In addition to the nearest neighbour in the manifold embedding y_p^i (blue dot in solid circle), the figure shows the other $\kappa - 1$ nearest neighbours (blue dots in dotted circles), the original point belonging to the manifold of \mathcal{S}_2 , i.e. y_q^j (red dot in yellow box), and the distances ω_i from which the similarities s_i are derived. The similarities are then used to form a weighted average of the corresponding motion fields.

motion field, i.e. a dense 3D motion field lacking the R-L component. This 3D motion estimate is the output of the motion model given the 2D surrogate image $b_p^{(new)}$ as input. Note that the coronal motion field from slice position \mathcal{C} is currently only used for the propagation of manifold coordinates but not for the reconstruction. This means the motion field from the coronal slice position will not be part of the volume.

Updating the Model and Adaptivity

The mechanism of embedding the new slice motion field (i.e. the motion model surrogate data) into the corresponding groups automatically updates the model. The new motion fields, after being used to stack a 3D motion field, stay in the model and may be used themselves in the future for new motion estimations.

In this manner, as the application phase goes on, more and more data is added to the model making it adaptive. In the case of respiratory drift or changes in the breathing pattern the model does not lose its validity but rather incorporates these new motion patterns.

6.3 Experiments and Results

In order to validate the proposed autoadaptive motion model (AAMM) technique it was compared to two versions of the method, each with one of the major novelties removed: AAMM without the autoadaptivity, and AAMM without the incorporation of slices of different orientations in the groupwise manifold alignment step. That is, the following techniques were compared:

- AAMM: The proposed autoadaptive motion modelling method as described in Section 6.2.
- AAMM (no adapt.): The proposed method without the adaptivity. This means that after each update step the most recently added 2D motion field was discarded and not incorporated into the model.
- SGA: The proposed AAMM method but without using the coronal input slices. Essentially, this is the SGA.REG technique extended to use sagittal motion fields instead of coronal images. The adaptivity was implemented in exactly the same way as for AAMM with the sole exception that there were no coronal input slices.

The experiments in this section aim to answer the following main research questions:

1. How does autoadaptivity affect the motion estimates after a short calibration phase with a constant breathing pattern?
2. Can the autoadaptive motion model adapt to a previously unseen breathing pattern?

In order to pursue these questions the three methods described above are evaluated on synthetic data derived from 6 volunteer scans and on real data acquired from 4 volunteers. In Experiment 1 (Section 6.3.2), synthetic data representing normal free breathing is generated to answer the first research question. In Experiment 2 (Section 6.3.3), additionally, synthetic data which corresponds to a deep breathing pattern is generated in order to investigate the second of the above questions. Lastly, in Experiment 3, the algorithms are evaluated on real volunteer scans acquired over 20 minutes. Using this data it is investigated how the methods respond to natural, gradual changes to breathing pattern which may not have been observed during model calibration.

Note that no comparison of AAMM to any other state-of-the-art motion modelling techniques was performed. All motion models from the literature follow the traditional motion modelling paradigm (see Fig. 6.1) and could not be built using the 2D slice-by-slice data used in this work. Thus an evaluation on equal terms was not feasible.

6.3.1 Parameter Choices

The free parameters of the investigated techniques were chosen based on the findings in Chapter 5 for SGA. Hence, here the following parameters were chosen for all of the methods: $\sigma = 0.5, \mu = 0.25, d = 3$. The parameter κ from Eq. (6.3), i.e. the number of nearest neighbours to use for the motion field estimation, was set to $\kappa = 3$.

The number of 2D motion fields increased steadily as the model was applied. Consequently, the neighbourhood size in the LLE cost function k was continually adapted to the

current data size. That is, after calibrating the model and upon adding an additional 2D motion field, k was set to

$$k = \text{round}(2\tau_p/3),$$

where τ_p is the number of 2D motion fields per slice position currently part of the respective groups.

6.3.2 Experiment 1: Synthetic Training Adaptivity

To quantitatively assess the proposed method in the presence of an approximately constant breathing pattern, very realistic synthetic 2D motion fields were generated by mimicking an actual slice-by-slice acquisition process. In this section, the generation was based on a 50 second dynamic 3D MR scan conducted under normal free breathing. The method for generating the data is described in detail in Appendix B.3. The final synthetic data is of the same form as the real data which results from the acquisition described in Section 6.2.1, and consists of 50 realistic motion fields per sagittal and coronal slice position. In addition, each of the slices is associated with a unique full 3D ground-truth motion field (including the S-I, A-P and R-L motion components). Note, that the resulting data mimics an acquisition of around 20-25 minutes. However, it can only reflect breathing patterns observed in the 50 second dynamic 3D MR scan. The extent of the respiratory motion was similar for the synthetic datasets of all volunteers. The average motion extent of the ground truth motion fields over all volunteers was 2.97 mm on average. The minimum average motion was 2.07 mm (volunteer C) and the maximum average motion was 3.71 mm (volunteer B). On average R-L motion accounted for 15.85% of the total motion, and the A-P and S-I accounted for 20.44% and 63.71%, respectively.

The motion estimation accuracy was quantitatively assessed using the three compared models on the synthetic slice-by-slice data. Each of the three stages of motion modelling shown in Fig. 6.1 was performed, i.e. model calibration, model formation and model application. The synthetic data generation as described in Appendix B.3 can be seen as a synthetic model calibration stage yielding slice-by-slice motion fields. In the next step, the model was formed by embedding a subset of the synthetic data using the three compared methods. 10 slices from each slice position were used for the initial formation of the model. Obtaining this amount of data in a real scan would take approximately 5 minutes. The motion model was then applied by continually adding all remaining slices one after the other, and at each time step evaluating the accuracy of the estimated motion against the 3D ground-truth motion field corresponding to the newest update slice.

In Fig. 6.6, the resulting motion estimation error curves are shown for all of the volunteers during a synthetic application phase. The evolution of the errors is shown over the duration of the application phase, which is the time it would take to acquire and add the

remaining slices in a real scenario. Here, an acquisition frequency of one slice per second was assumed which corresponds to a heart rate of 60 beats per minute. Each point in Fig. 6.6 represents the mean error obtained over a time interval of $2L$ update slices, where L is the number of slice positions in the region of interest. That means the time interval is equal to the time taken to acquire each slice position twice.

In order to quantitatively evaluate the 3D motion estimation errors and the adaptivity of the compared techniques, the application phase was split into 5 time periods T_1, \dots, T_5 , of equal length. Those are highlighted in Fig. 6.6. In Table 6.1 the mean 3D motion estimation errors in the corresponding time intervals are summarised for the three methods and for all 6 volunteers.

Table 6.1: Results of synthetic validation on normal breathing data. The table shows average 3D motion estimation errors in mm obtained by each of the investigated techniques compared to the ground-truth 3D motion fields.

Volunteer	Method	T_1	T_2	T_3	T_4	T_5
		<i>mm</i>	<i>mm</i>	<i>mm</i>	<i>mm</i>	<i>mm</i>
A	SGA	1.97	1.68	1.59	1.56	1.52
	AAMM (no adapt.)	1.65	1.66	1.68	1.70	1.74
	AAMM	1.51	1.38	1.37	1.32	1.34
B	SGA	2.34	2.01	1.91	1.88	1.68
	AAMM (no adapt.)	2.03	1.99	2.06	2.08	1.94
	AAMM	1.90	1.66	1.63	1.59	1.40
C	SGA	1.43	1.33	1.22	1.37	1.15
	AAMM (no adapt.)	1.28	1.35	1.20	1.37	1.26
	AAMM	1.25	1.18	1.05	1.17	1.05
D	SGA	1.80	1.38	1.33	1.11	1.18
	AAMM (no adapt.)	1.56	1.39	1.45	1.39	1.52
	AAMM	1.42	1.11	1.15	1.04	1.11
E	SGA	1.89	1.67	1.58	1.44	1.28
	AAMM (no adapt.)	1.72	1.73	1.68	1.78	1.68
	AAMM	1.60	1.38	1.33	1.21	1.18
F	SGA	2.20	1.75	1.74	1.63	1.59
	AAMM (no adapt.)	2.05	1.73	1.90	1.79	1.93
	AAMM	1.79	1.47	1.51	1.35	1.44

For all volunteers the AAMM technique significantly outperformed the other two methods in all of the intervals as can be seen in Table 6.1, and also by comparing to the error curves shown in Fig. 6.6. Significance was assessed using a 1-tailed Wilcoxon signed rank test as in the previous Chapter. The estimation errors for AAMM and its non-adaptive counterpart, AAMM (no adapt.), were similar in the beginning of the application phase, but as anticipated, as the application phase went on, the AAMM technique continually improved its accuracy by incorporating more and more data into the model. On average

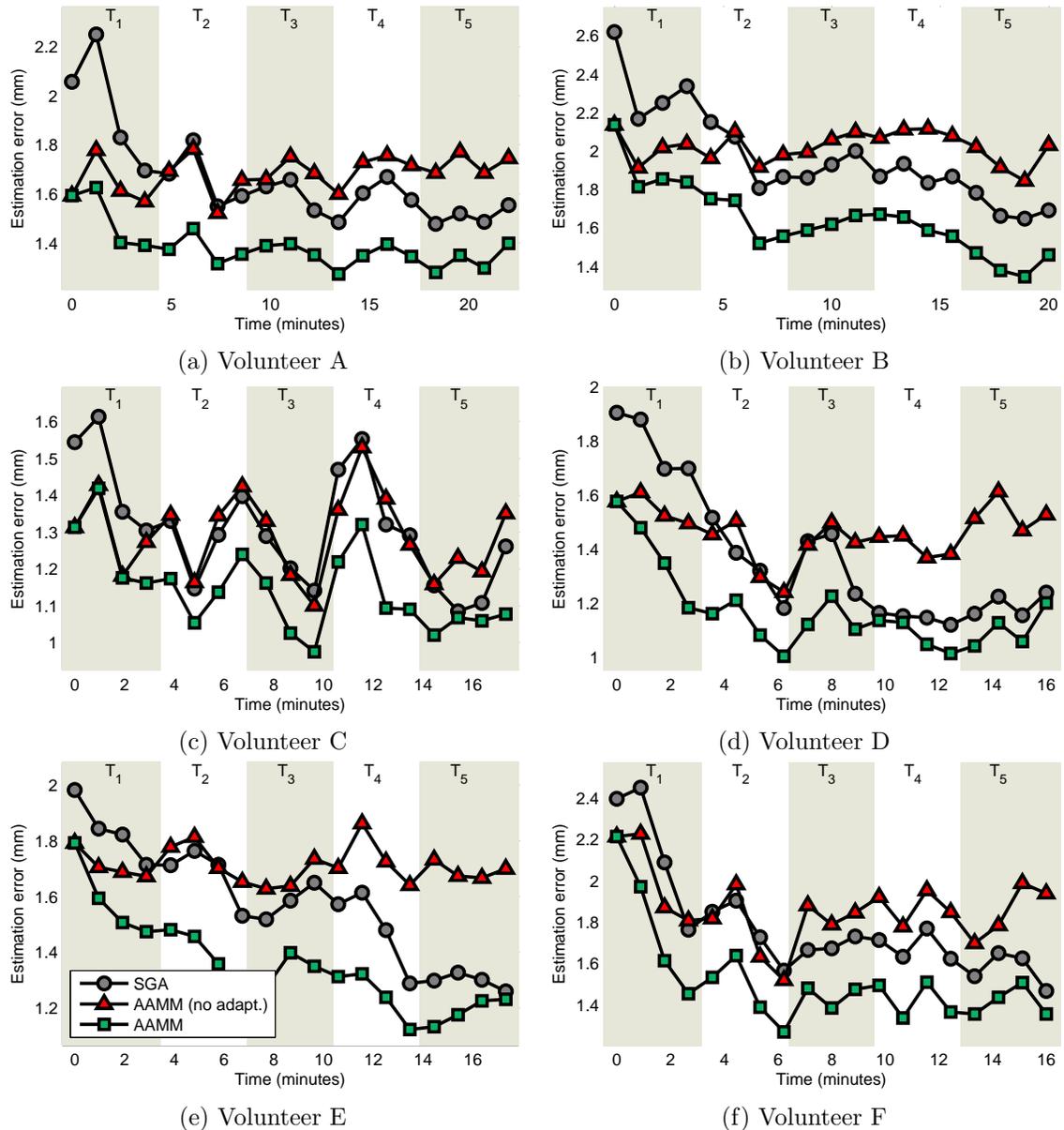


Figure 6.6: Average 3D motion estimation errors in mm for all volunteers over the entire duration of the synthetic application phase for normal breathing data.

the motion estimation of AAMM improved by 22.94% in T_5 with respect to its non-adaptive counterpart. However, significant improvements of 16.87% on average were already reached in T_2 , i.e. after approximately 3 to 7 minutes of treatment. By visually inspecting the curves for AAMM in Fig. 6.6 it can be seen that for many of volunteers (in particular volunteers A, D, E, and F) the error curves start to saturate approximately around the 7 minute mark. From this it can be concluded that a longer calibration scan of around

12 minutes would be optimal, that is the 5 minutes that were used for calibration in this experiment plus 7 minutes worth of data added during the application phase.

The AAMM technique also consistently performed better than SGA, i.e. the version without coronal slices. This shows that the addition of data from a coronal slice position in the manifold alignment step improves the 3D motion estimation accuracy.

A fraction of the remaining errors was due to the fact that the technique currently cannot estimate R-L motion. In this experiment on synthetic data the R-L motion was responsible for on average 46.53% of the error of the AAMM technique, or on average 0.63 mm.

Note that the error curves did not necessarily steadily decrease over the entire period of time, but exhibited some variations usually affecting all methods equally. See for example T_3 and T_4 of volunteer C (Fig. 6.6c). The motion estimation error tends to be smaller for exhale motion states than for inhale motion states, since the motions involved are smaller. The variations in the error can be explained by differences in the frequency of occurrence of exhale or inhale states. For example, there were a large amount of inhale motion states around the 11 minute mark of volunteer C, and a large amount of exhale states around the 9 minute mark. The variations in the error between volunteers can be explained by the fact that the synthetic data was derived from real volunteer scans and some volunteers naturally had larger or more complicated motion patterns.

6.3.3 Experiment 2: Synthetic Adaptivity to New Breathing Pattern

In Experiment 1, it was investigated how the autoadaptive technique behaves if more data of the same breathing pattern is added. However, the data used in that experiment does not reflect any of the long term changes which may occur in real data, such as drift or changes in breathing mode. In order to investigate if the model can adapt to previously unseen breathing patterns, for this experiment, a second synthetic dataset was generated using the same methodology as above (described in Appendix B.3), but using a 50 second dynamic 3D MR scan performed under deep breathing as input. That scan was performed immediately after the 50 second free breathing scan, but the volunteers were instructed to take deep quiet breaths. For all volunteers this resulted in synthetic data with significantly longer respiratory cycles and significantly larger displacements of the anatomy. The average magnitude of ground truth motion over all volunteers was 9.22 mm. However, the magnitude of the displacements varied significantly from volunteer to volunteer with a minimum of 4.59 mm (volunteer C) and a maximum of 13.48 mm (volunteer F). On average R-L motion accounted for 17.59% of the total motion, and the A-P and S-I accounted for 23.76% and 58.65%, respectively.

In order to investigate how the examined methods would react to this new deep breath-

ing pattern, in a first step the models were calibrated and formed by using all time points of the normal breathing data. This means that the models had largely adapted to the normal breathing pattern. Note that the state of the models was the same as for the last time point of the AAMM technique in Figure 6.6. In a next step, the motion models were applied using the synthetic deep breathing data. That is, the 2D deep breathing motion fields were added to the model one-by-one, and the motion estimation error was evaluated exactly as in Experiment 1. The resulting error curves are shown in Figure 6.7, where each point corresponds to an average over $2L$ motion estimates. In order to assess the performance of the models quantitatively, the errors for each subject were averaged within 5 time intervals of equal length. Those errors are presented in Table 6.2 for all volunteers.

Table 6.2: Results of synthetic validation on deep breathing data. Average 3D motion estimation errors in mm obtained with each of the investigated techniques compared to the ground-truth 3D motion fields.

Volunteer	Method	T_1	T_2	T_3	T_4	T_5
		<i>mm</i>	<i>mm</i>	<i>mm</i>	<i>mm</i>	<i>mm</i>
A	SGA	4.23	4.01	3.86	4.00	4.07
	AAMM (no adapt.)	4.43	4.51	4.48	4.48	4.56
	AAMM	3.90	3.73	3.53	3.55	3.65
B	SGA	6.82	6.32	6.41	6.59	6.51
	AAMM (no adapt.)	7.47	7.35	7.70	7.69	7.46
	AAMM	6.32	5.55	5.67	5.82	6.32
C	SGA	3.15	2.88	2.86	2.74	2.61
	AAMM (no adapt.)	3.09	3.10	3.23	3.13	3.08
	AAMM	2.78	2.52	2.55	2.38	2.31
D	SGA	9.69	8.89	8.77	7.88	8.62
	AAMM (no adapt.)	11.61	11.85	11.75	10.15	11.94
	AAMM	8.84	7.56	7.57	6.54	7.56
E	SGA	6.33	5.88	5.72	5.45	5.59
	AAMM (no adapt.)	6.26	6.44	6.54	6.51	6.32
	AAMM	5.83	5.16	4.94	4.66	4.61
F	SGA	9.94	8.24	8.57	8.73	8.91
	AAMM (no adapt.)	11.34	9.67	11.32	11.41	11.71
	AAMM	9.17	7.33	8.21	8.05	8.13

As before the AAMM method significantly outperformed the two other techniques for all volunteers and for all time intervals as can be seen by comparing the error curves in Figure 6.7, and the error figures in Table 6.2. As expected, the estimation errors for AAMM and its non-adaptive counterpart AAMM (no adapt.) started at similar values in T_1 , but AAMM led to improved motion estimates the more data of the new breathing type was added to the model. Already in time-interval T_2 , AAMM led to significant average improvements of 21.45% over AAMM (no adapt.) In T_5 , the average improvements amounted to 27.10%. As

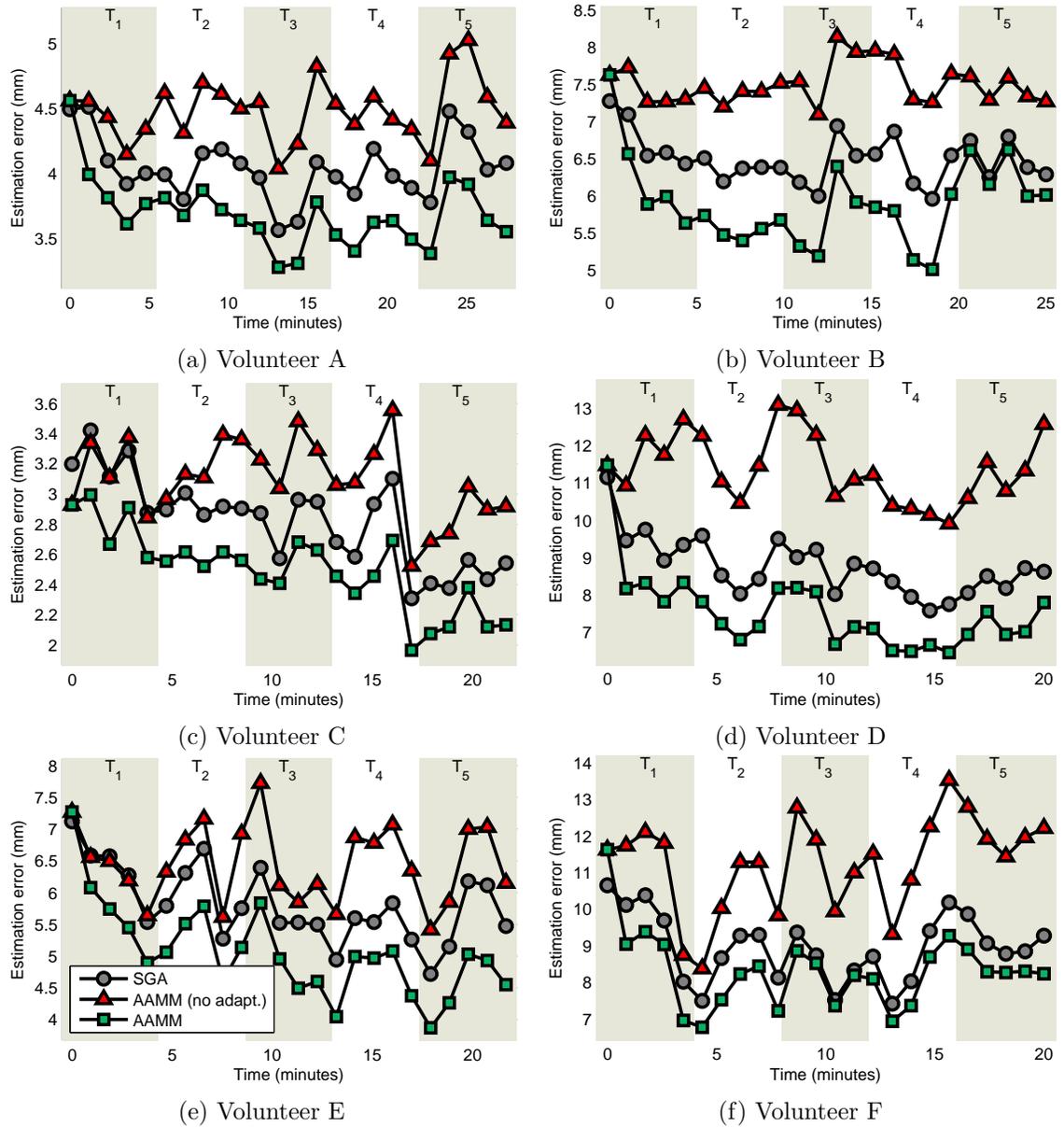


Figure 6.7: Average 3D motion estimation errors in mm for all volunteers over the entire duration of the synthetic application phase when the breathing type was changed to deep breathing. The entire interval shown contains only deep breathing.

before AAMM also performed significantly better than the SGA technique, which is due to the additional robustness added by the coronal slice employed in the AAMM technique.

For all examined methods the motion estimation errors were significantly larger for the deep breathing pattern than for the normal breathing pattern in Experiment 1. This is due to the fact that the deep breathing data contained much larger motion amplitudes. The variations between the subjects are due to the fact that the extent of the deep breathing

motion varied from volunteer to volunteer. The average error over all subjects due to the missing motion estimates in the R-L direction amounted to 42.15% of the motion estimation error of AAMM, or 2.28 mm.

6.3.4 Experiment 3: Adaptivity on Real Data

For the experiments on real data real dynamic slice-by-slice data and a slice-by-slice breath-hold volume were acquired as described in Section 6.2.1. In order to validate the model, the data for the calibration and model formation, and for the model application were acquired in one long scan. Overall, each slice position was acquired 40 times which typically resulted in an approximately 20 minute scan. Additionally, a 1D pencil beam navigator signal was acquired from the left hemi-diaphragm immediately before the acquisition of each 2D slice, which was used to validate the accuracy of the motion estimations, but was not used not for any part of the motion modelling framework.

As in Experiment 1, the three models were formed on the motion fields derived from the first 10 slices acquired from each slice position, which corresponds to a calibration phase of 5 minutes. During the model application phase, the remainder of the slices were added one by one and a 3D motion field was estimated for each of the input slices. Because for the real data no ground-truth motion was available, instead, the slice-by-slice breath-hold volume was transformed using each estimated 3D motion field and a 1D navigator value was extracted from a rectangular region of interest on the dome of the left hemi-diaphragm, i.e. approximately the same location from which the real pencil beam navigator was acquired. Given a perfect motion estimation this navigator signal should be strongly correlated with the pencil beam navigator. In reality, however, this correlation depends on the accuracy of the estimated 3D motion.

In order to assess the adaptivity of the compared methods, the normalised cross correlation (NCC) of the estimated navigator signal with the pencil beam navigator was measured over fixed intervals. In Fig. 6.8 the progression of this correlation for all four volunteers is shown. For a robust estimation of the NCC, time intervals of $2L$ were chosen to calculate each error point.

As for the synthetic data the entire application phase was divided into 5 larger time intervals T_1, \dots, T_5 . In Table 6.3 the NCC between the pencil beam navigator and the retrospectively derived navigator signal is reported for all volunteers over the entire duration of these periods.

AAMM outperformed the other two methods for most time intervals. Furthermore, as for the synthetic data, it could again be observed that the motion estimation accuracy, as measured by NCC, improved over the duration of the application phase. Note that the data in this experiment was derived from a relatively long scan, where natural changes in

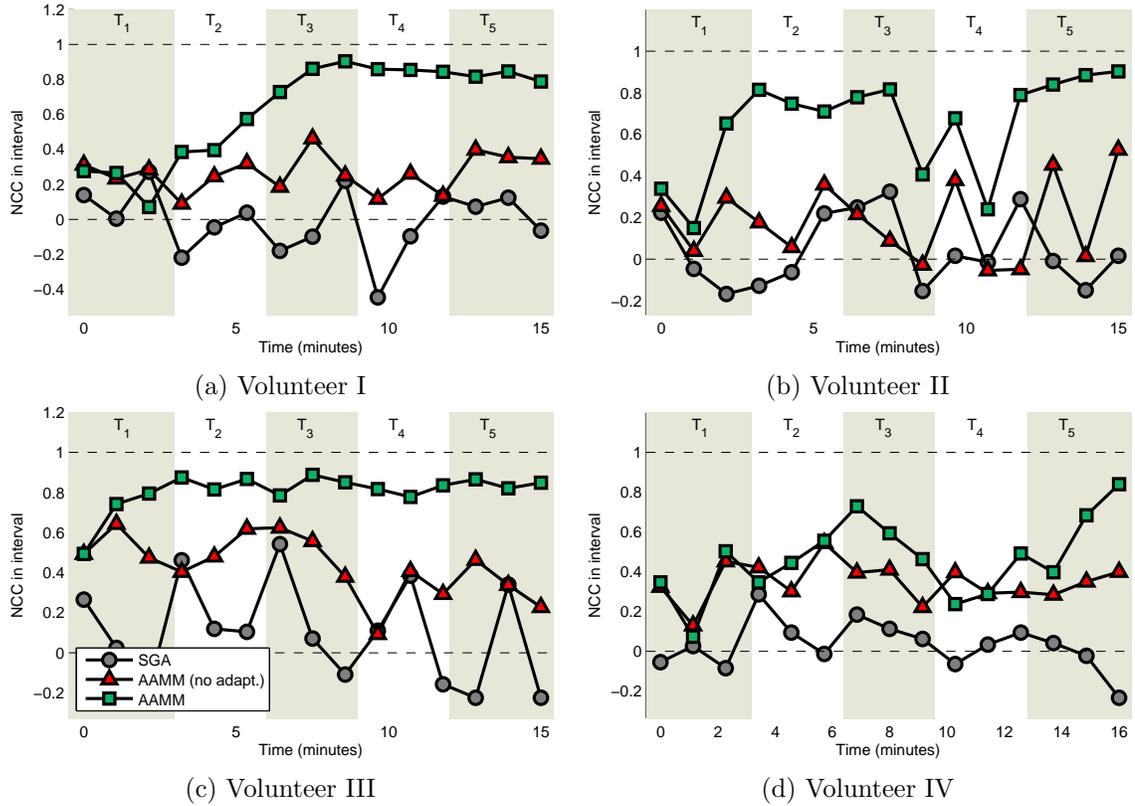


Figure 6.8: NCC of pencil beam navigator with a navigator signal derived from a volume deformed using a 3D motion estimation provided by each of the three investigated methods over time.

Table 6.3: NCC of a pencil beam navigator and a navigator signal derived from a volume deformed using a 3D motion estimation provided by each of the three investigated methods. The results are shown over the five time intervals indicated in Fig. 6.8.

Volunteer	Method	T_1	T_2	T_3	T_4	T_5
I	SGA	0.15	-0.08	-0.08	-0.06	0.05
	AAMM (no adapt.)	0.26	0.25	0.29	0.24	0.35
	AAMM	0.19	0.49	0.83	0.85	0.83
II	SGA	-0.01	-0.07	0.06	0.04	-0.06
	AAMM (no adapt.)	0.16	0.17	0.09	0.05	0.42
	AAMM	0.36	0.75	0.56	0.41	0.86
III	SGA	0.05	0.19	0.18	0.08	-0.10
	AAMM (no adapt.)	0.55	0.54	0.50	0.32	0.31
	AAMM	0.67	0.85	0.85	0.81	0.84
IV	SGA	-0.01	0.09	0.04	0.07	-0.09
	AAMM (no adapt.)	0.32	0.45	0.31	0.32	0.35
	AAMM	0.33	0.48	0.55	0.40	0.69

respiration patterns are very likely to happen due to relaxation or, occasionally, due to the volunteer falling asleep in the scanner. It can be observed that motion estimation accuracy sometimes dropped due to such changes. For example, volunteer II (see Fig. 6.8b) started taking deep breaths around T_3 , but then returned to his previous breathing pattern. Had he continued breathing deeply presumably the model would have adapted to that pattern. Note that for volunteer III the NCC quickly approaches its maximum for AAMM, but continually decreases for its non-adaptive counterpart. By examining the original pencil beam navigator signal shown in the top row of Fig. 6.9 it can be seen that the subject exhibited a significant drift in their respiration base level throughout the imaging session. By comparing this signal to the signals estimated by AAMM and AAMM (no adapt.) it can be observed that AAMM manages to follow this drift whilst AAMM (no adapt.) cannot adapt its range of motion estimations.

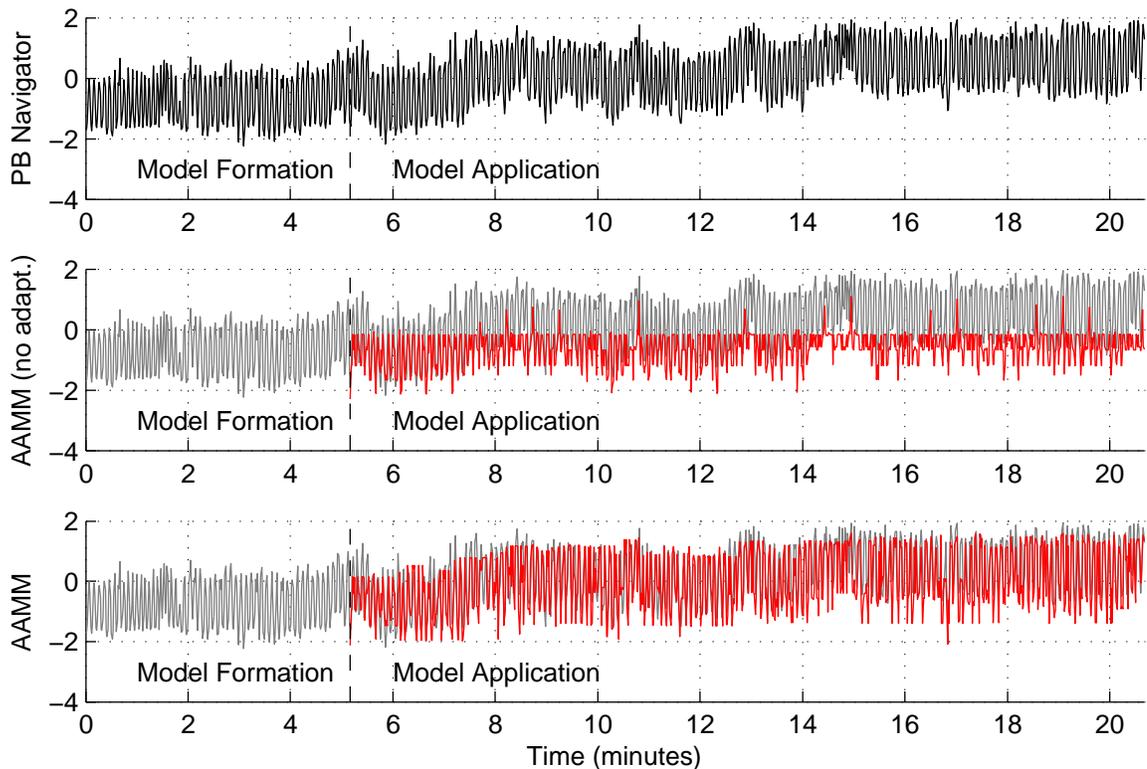


Figure 6.9: Example of the pencil beam navigator signal acquired for validation for volunteer III (top row) and navigator estimations produced by AAMM (no adapt.) (middle row) and AAMM (bottom row). The original pencil beam navigator is underlaid in grey for comparison. The entire time interval is shown including the approximately 5 minutes of the calibration scan, and the approximately 15 minutes of model application. All signals have been normalised using the mean and standard deviation such that the pencil beam navigator signal and the estimated signals can be directly compared.

As opposed to the synthetic experiments, here SGA consistently performed worse than the other examined methods. It was found that on real data SGA was fundamentally not robust to the sagittal input slices. It was observed that respiratory information often failed to propagate through the body middle, that is for example, an input slice from the left body half would often fail to properly estimate motion in the right body half and vice versa. Incorporating data from coronal slice positions in the manifold alignment step of AAMM effectively solved this problem.

6.4 Discussion

In this chapter, a novel motion modelling paradigm which enables accurate 3D motion estimations over extended periods of time in the scenario of MR-guided interventions was proposed. This is achieved by using partial motion information, i.e. 2D motion fields estimated from MR slices, to form as well as to apply the motion model. The fact that the calibration and surrogate data are of the same type inherently enables the proposed motion model to automatically adapt to changing breathing patterns without the need to rebuild the model during the application phase.

The vast majority of motion models in the literature cannot adapt to changing breathing patterns and need to be rebuilt entirely if the correlation between the surrogate and the motion data loses validity [McClelland et al., 2013]. A small number of papers such as Schweikard et al. [2005] and King et al. [2012] proposed adaptive techniques which have the ability to intermittently acquire small amounts of new calibration data during a treatment and rebuild the model intra-procedure. However, none of the existing motion models in the literature are able to continually adapt to changing patterns in the whole thorax such as the proposed AAMM technique.

The autoadaptive motion model is implemented by extending the SGA technique proposed in Chapter 5 to use 2D motion fields as input. This approach has two important limitations: Through-plane motion may distort the motion estimations and motion in the direction orthogonal to the slices cannot be estimated. Suh et al. [2008] have shown that lung tumour motion due to respiration is largest in the S-I or A-P directions for the majority of patients. Park et al. [2012] found that liver tumour motion is smallest in the R-L direction with a magnitude of 3.0 mm on average. In comparison, the average motions in the S-I and A-P directions amount to 17.9 mm and 5.1 mm, respectively. Hence, in order to minimise the effects of through-plane motion, the motion estimates were derived from sagittal input slices. The large differences in the appearance of sagittal slices acquired from different locations necessitated the incorporation of coronal images from a single slice position. In order to combine sagittal and coronal data the methodology of SGA was substantially extended to arrive at the proposed AAMM technique.

The proposed method was validated on realistic synthetic and real data. The experiments show that the autoadaptive motion model is able to improve its motion estimations over time and can thus produce significantly better 3D motion estimations over the duration of an MR-guided treatment compared to non-adaptive methods. Furthermore, the experiments show that the incorporation of data from a single coronal slice position leads to significant improvements in motion estimation. Note that the performance of AAMM was not compared to traditional motion models from the literature. The presented method follows an entirely novel motion modelling paradigm estimating 3D motion from partial motion estimates and uses 2D MR data in all stages of the model. This is conceptually different from the classical surrogate-driven motion model paradigm where typically 3D MR data would be acquired prior to the treatment along with some simpler surrogate signal. Hence, it was not possible to compare against existing techniques using the same data.

The proposed method offers a new way of performing motion modelling. However, in its current form it still suffers from some drawbacks. Currently the motion is estimated only from sagittal 2D MR slices. This orientation was chosen to minimise the effects of through-plane motion. However, the remaining R-L motion may cause artefacts in the registration step of the calibration phase. Furthermore, the 3D motion estimations are obtained by simply stacking the 2D motion fields, which leads to motion fields lacking the R-L component. In the evaluations on synthetic data the missing R-L component accounted for over 40% of the remaining motion estimation error. Note that currently the coronal motion fields are used only in the groupwise manifold alignment step but are not part of the final 3D motion estimations. It may be possible to mitigate the through-plane motion effects by also using motion information derived from one or potentially several coronal slice positions in the 3D motion estimation step.

Another drawback of the current implementation is that it suffers from significant latency. Given the current image acquisition and computation speeds of the non-optimised code, the motion modelling system would have a latency of around 800 ms. One option to circumvent this problem could be to combine the method with a motion prediction technique such as the one proposed by Sharp et al. [2004]. However, to a certain degree the problems of acquisition speed and latency can be addressed directly.

In the present work cardiac gated images were employed in order to isolate the respiratory motion. This effectively limits the update frequency of the model to approximately 1 Hz based on a typical heart rate of 60 beats per minute. The cardiac gating, however, is not an essential part of the technique and could be easily dropped if the region of interest excluded the heart as in the scenario of a MR-guided HIFU of the liver. 2D MR images can be acquired in less than 50 ms [King et al., 2012], and potentially even as fast as 20 ms [Uecker et al., 2010] by making use of parallel imaging techniques. Furthermore, it

has been shown that both 2D MR image reconstruction and 2D image registrations can be performed in less than 20 ms by employing GPU implementations [Hansen et al., 2008; Kubias et al., 2008]. Lastly, significant speed improvements could also be achieved for the groupwise manifold alignment with a more efficient parallel implementation. It is therefore believed that it may be possible to run an autoadaptive motion modelling system close to real-time with update times of less than 100 ms. In the above experiments it was found that the proposed system can significantly improve the motion estimation accuracy by close to 20% in less than 10 minutes. Potentially, however, much faster adaptivity could be achieved. Based on a hypothetical update frequency of 100 ms, significant improvements in the motion estimation in the whole thorax could be achieved in as little as 1 minute.

Lastly, in the autoadaptive motion model in its present form all the data added is retained. The rationale behind this is that, in this manner, the model can go back to breathing patterns which were observed before a change occurred. A patient may for example go back and forth between a calm and a nervous breathing pattern as a result of certain actions of the surgeon or the progress of the treatment. However, the larger the model grows the more memory is used to store the 2D motion fields and the more computationally expensive it becomes to evaluate the updated group embeddings. It may therefore make sense to implement a “ring buffer” approach, where older data is discarded as new data is added to the model. An interesting future direction would be to automatically determine which data is essential to model certain breathing types and selectively delete data which is unlikely to be used again.

Chapter 7

Self-Aligning Manifolds

Contents

7.1	Introduction	115
7.2	Method	117
7.2.1	Similarity Without Correspondences	117
7.2.2	Simultaneous Embedding of Many Datasets	119
7.3	Applications	120
7.3.1	Application to 4D MR Reconstruction	120
7.3.2	Application to 3D Ultrasound Gating and Compounding	121
7.4	Experiments and Results	121
7.4.1	Experiment 1: Synthetic 4D MR Reconstruction	121
7.4.2	Experiment 2: 4D MR Reconstruction from Real Data	122
7.4.3	Experiment 3: 3D Ultrasound Compounding	123
7.5	Discussion	123

7.1 Introduction

In the previous two chapters the objective was to match the respiratory positions of multiple datasets comprised of 2D MR slices, or motion fields derived from such slices, in a common low-dimensional space. In order to derive the similarity kernel for the simultaneous embedding, high-dimensional data had to be directly compared. Since only a subset of the datasets were comparable in high-dimensional space, e.g. only data from neighbouring slice positions, the data was embedded in groups of two datasets each

This restriction significantly limits the applicability of the method. As was discussed in Chapter 6, matching sagittal slices in this manner may be inaccurate even for data

from neighbouring slice positions and respiratory information may get lost when propagating through the body middle. In some manifold alignment applications comparison in the high-dimensional space may not be possible at all due to the different appearance of the data or potentially due to the datasets being of a different type altogether. This may, for example, be the matching of images of a population acquired with different imaging protocols [Guerrero et al., 2014], matching respiratory motion states in different modalities [Thielemans et al., 2011] or matching motion states in dynamic images of the same modality acquired from different views. An example of the latter is the compounding of free breathing 3D ultrasound images from different views for the purpose of reducing artefacts and increasing the effective field of view [Yao et al., 2011]. Thus it would be desirable to perform a manifold alignment without the need of any comparisons in the high-dimensional space.

A small number of papers in the related literature have proposed manifold alignment techniques which require neither prior correspondences nor comparability of the high-dimensional data. Those include the two-step approaches by Mateus et al. [2008, 2007] for articulated shape matching of human figures, and by Pei et al. [2012] for non-medical image matching. The most closely related work was that by Wang and Mahadevan [2009], who proposed to establish correspondences between document collections using a one-step manifold alignment approach by considering similarities of the local neighbourhood within each of the datasets to be aligned. Unfortunately, the approach requires iterating through all of the permutations of the local neighbourhood of each point, which is only computationally tractable for very small neighbourhoods. The only work from the medical imaging literature which can be considered unsupervised is the modality independent feature descriptor for image registration by Wachinger and Navab [2012, 2010]. However, in that work only 1D embeddings are considered and alignment is performed by a simple scaling step. Such an approach would not be possible for two-, or higher-dimensional embeddings.

In this chapter, a robust one-step manifold alignment method for the ‘self-alignment of manifolds’ (SAM) of medical images is proposed. In the following, LLE is extended in a manner which allows the manifolds of multiple datasets to be mapped into a globally consistent space without prior correspondences or inter-dataset comparisons in the high-dimensional space. The proposed method can be used to align manifolds constructed from medical imaging datasets which depict different anatomy or the same anatomy from different views, and that are hence not directly comparable. The technique is inspired by work from the graph matching literature, and establishes correspondences between multiple datasets based on their internal graph structure.

The method is demonstrated on two problems in medical imaging which require matching of respiratory positions of different views which are hard to compare in image space: the

4D reconstruction of MR volumes from sagittal slices acquired at different slice positions and the compounding of free breathing 3D US sequences acquired from two different views.

This work has been presented in Baumgartner et al. [2015].

7.2 Method

In the previous two chapters groupwise embeddings were employed. However, as was shown in Chapter 6, when using groupwise embeddings, the matching may fail completely if applied to sagittal MR slices without using coronal slices for guidance. The main reason for that is that respiratory information cannot successfully propagate through the body middle. Even in other regions, and also for coronal input slices, the groupwise methodology is prone to error propagation. A one-step approach would not require any propagation because all the points are embedded in one single, globally consistent space. However, in Chapter 5 it was found that extending LLE to more than two datasets did not yield good reconstructions if applied using an image-based similarity kernel. It was furthermore found in preliminary experiments that leaving the kernels U_{pq} undefined for $|p-q| > 1$ did not lead to satisfactory alignment of the datasets.

In this chapter, a novel method for determining the similarity kernel U_{pq} which allows computation of a similarity between datasets X_p, X_q without having to relate them in the high-dimensional space is introduced. Using this novel kernel, the one-step approach from Chapter 5 is revisited to create a robust manifold alignment technique, which does not require the input datasets to be comparable in the high-dimensional space.

In the following, first the novel similarity kernel will be described in Section 7.2.1 and then, in Section 7.2.2, it will be shown how it can be used in a one-step approach.

7.2.1 Similarity Without Correspondences

Most work on one-step manifold alignment has either used known prior correspondences to define the similarity kernels U_{pq} [Bhatia et al., 2012a; Guerrero et al., 2014], or has used similarity measures between the high-dimensional data to define it [Lombaert et al., 2013; Torki et al., 2010]. For many applications, however, it is desirable to simultaneously embed medical image datasets which are not comparable in image space and for which prior correspondences are not easily obtainable. This problem is effectively one of finding a suitable inter-dataset similarity kernel U_{pq} that connects datasets X_p, X_q but which is *not* based on comparisons between datasets in the high-dimensional space.

In this chapter a novel, robust method for deriving such an inter-dataset similarity kernel is proposed. Instead of directly comparing the data from different datasets in the high-dimensional space, first the internal graph structure of each dataset is analysed and a

feature vector characteristic of the local neighbourhood of each node is derived. In the next step these feature vectors are compared across datasets. This approach is an extension of the method proposed in Gori et al. [2005] for graph matching.

Deriving a Characteristic Feature Vector for Each Node

For each dataset X_p a fully connected, weighted graph $\mathcal{G} = \{V, E, C\}$ is formed which connects each node in V to every other with edges E . The edge weights C^{ij} connecting node V^i to V^j are given by a Gaussian kernel

$$C^{ij} = e^{-\frac{\|x_p^i - x_p^j\|^2}{2\sigma_1^2}},$$

where x_p^i, x_p^j are the high-dimensional data points of X_p , and σ_1 is a shape parameter. Note that C^{ij} is formed by intra-dataset comparisons only.

As noted in Gori et al. [2005], the steady-state distribution of a (lazy) random walk, i.e. the probability \mathbf{p}_i of ending up at a particular node V^i from each other node after $t \rightarrow \infty$ time steps, can provide information about a node's location in the graph. In a lazy random walk the walker remains at their current node V^i with probability 0.5 at each time step and walks to a random neighbour V^j with a probability given by the edge weights C^{ij} the other half of the time. The state of the random walk in a weighted graph after $t + 1$ time steps is given by

$$\mathbf{p}^{(t+1)} = \frac{1}{2}(I + CB^{-1})\mathbf{p}^{(t)}, \quad (7.1)$$

where B is the diagonal degree matrix given by $B^{ii} = \sum_j C^{ij}$. The steady state distribution is stationary and is given by

$$\mathbf{p}^* = \frac{B\mathbf{1}}{\mathbf{1}^T B\mathbf{1}}. \quad (7.2)$$

In Gori et al. [2005], \mathbf{p}^* was used to perform graph matching. However, in preliminary experiments, \mathbf{p}^* alone was not sufficiently robust to align the datasets. Therefore, to increase the robustness of this measure, here additionally the neighbourhood of each node is analysed. That is, for each V_i the r -nearest neighbours are identified and their \mathbf{p}^* are evaluated as well, by forming an average over the neighbourhood:

$$\pi_i^{(r)} = \frac{1}{|\eta(i)|} \sum_{j \in \eta(i)} \mathbf{p}_j^*. \quad (7.3)$$

In evaluations it was found that using a single value of r does not produce a rich enough descriptor of local graph structure. Therefore, instead the value of r is systematically

increased from 1 to a maximum neighbourhood size R , resulting in a feature vector

$$\mathbf{r}_i = [\mathbf{p}_i^*, \pi_i^{(1)}, \dots, \pi_i^{(R)}] \quad (7.4)$$

for each node V^i . This feature vector is highly characteristic for each node in the graph for large enough values of R . As an example one can imagine the case of the nodes corresponding to medical images at different respiratory states. Because of the reproducibility of exhales and the fact that most people spend more time at exhale than inhale, an exhale image will have many close neighbours with similar \mathbf{p}^* values but only few far away neighbours with varying \mathbf{p}^* values, whereas an inhale image will exhibit the opposite behaviour.

Matching Graph Nodes Across Datasets

The feature vectors $\mathbf{r}_i^{(\ell)}$ are derived for each dataset ℓ as described above. Next, a similarity kernel \tilde{U}_{pq} , connecting two datasets p and q , is formed as

$$\tilde{U}_{pq}^{ij} = e^{\frac{-\|\mathbf{r}_p^i - \mathbf{r}_q^j\|^2}{2\sigma_2^2}}, \quad (7.5)$$

where σ_2 is another shape parameter. Lastly, a one-to-one matching between the two datasets found using the Hungarian algorithm as was done before in Chapters 5 and 6, and also in Gori et al. [2005]. This forms the final similarity kernel U_{pq}^{ij} , which can be substituted into Eq. (5.1) from Chapter 5. Note that the non-zero coefficients of U_{pq} still contain the original similarities from Eq. (7.5) such that a very good match will have a higher impact on the alignment than a moderate match.

7.2.2 Simultaneous Embedding of Many Datasets

The extension of the one-step approach to L datasets was already discussed previously in Section 5.2.2. It will be partially reproduced here for the reader's convenience.

Analogous to the case with only two datasets covered in the Theory chapter (Chapter 4), the total cost function C_{tot} which needs to be minimised can be rewritten in matrix form as

$$C_{tot} = Tr(VHV^T), \quad (7.6)$$

where V is the matrix containing the concatenated embeddings, $V = [Y_1, \dots, Y_L]$, and

$$H = \begin{bmatrix} M_1 + \mu \sum_p D_{1p} & -\mu U_{12} & \dots & -\mu U_{1L} \\ -\mu U_{21} & M_2 + \mu \sum_p D_{2p} & \dots & -\mu U_{2L} \\ \vdots & & \ddots & \vdots \\ -\mu U_{L1} & -\mu U_{L2} & \dots & M_L + \mu \sum_p D_{Lp} \end{bmatrix}. \quad (7.7)$$

Here, the diagonal degree matrices D_{pq} are given by $D_{pq}^i = \sum_j U_{pq}^{ij}$. Under the constraints discussed in Chapter 5, and using the novel similarity kernel defined in Eq. (7.5), the embeddings V are given by the second smallest to $d + 1$ smallest eigenvectors of H .

7.3 Applications

The novel self-aligning manifolds algorithm is applied to two problems in medical image analysis: 1) the reconstruction of accurate 4D MR sequences of the liver from sagittal 2D slices, and 2) the simultaneous gating and compounding of 3D liver US images.

7.3.1 Application to 4D MR Reconstruction

High-resolution 2D MR slices were acquired from L anatomical slice positions such that the whole liver was covered with several acquisitions per slice position. As in Chapter 6 sagittal slices were used because most respiratory motion occurs in the sagittal plane. The acquisitions were carried out on a Philips Achieva 3T MR scanner using a T1-weighted gradient echo sequence with an acquired in-plane image resolution of $1.4 \times 1.4 \text{ mm}^2$ and a slice thickness of 8 mm. The slice positions were not overlapping in this work, in order to decrease overall acquisition times. To cover the liver, typically around 30 slice positions were needed and each slice position was sampled 100 times. Furthermore, here no cardiac gating was used because the effects of cardiac motion are negligible in the liver. Since a single slice could be acquired in $\sim 270 \text{ ms}$ the typical total acquisition time amounted to ~ 14 minutes. For the synthetic experiments such data is generated in a realistic manner from other scans. The process is similar to the one used in Chapter 6 and is described in detail in Appendix B.3.

The image data from the L slice positions form L high-dimensional datasets X_1, \dots, X_L , where each X_p contains all slices acquired at slice position \mathcal{S}_p . As before in Chapter 5, for each of the acquired slices the aim is to reconstruct a 3D volume by stacking appropriate slices from all other slice positions such that the respiratory position is consistent in the whole volume. The sequence of reconstructed 3D volumes from each slice results in the 4D MR sequence. To achieve this the novel self-aligning manifolds (SAM) technique is applied to reduce the dimensionality of these datasets and embed them in a globally consistent way (i.e. align the manifolds). In this manner, the proximity of two low dimensional points y_p^i, y_q^j from slice positions \mathcal{S}_p and \mathcal{S}_q , implies proximity (in respiratory position) of the high-dimensional images x_p^i, x_q^j . Thus, for the 4D MR reconstruction it suffices to look up the nearest neighbour of each slice in the low-dimensional space for all other slice positions and stack the corresponding high-dimensional slices into a volume.

7.3.2 Application to 3D Ultrasound Gating and Compounding

For this application, freehand 3D US images of the liver from two overlapping views were acquired for 22 seconds at 14 frames/second (i.e. ~ 300 frames per view). A Phillips iE33 US system with a X3-1 matrix array transducer was used. The data from the two views make up two datasets X_1, X_2 . The first 2 seconds of the acquisition from each view were performed at an end-exhale breath-hold.

The compounding of the two views leads to an extended field of view and reduced noise artefacts. The objective is to obtain a number of such compounded volumes, each of which has a consistent respiratory position. For this purpose the imaging data from both views are binned in the low-dimensional space such that all images in one bin originate from approximately the same respiratory state. Next, for each bin a compounded volume is reconstructed.

The dimensionality of the two US sequences X_1, X_2 is reduced using SAM such that their embeddings Y_1, Y_2 lie in a globally consistent space. Although higher dimensions d of the embedding will be able to capture more respiratory variability the dimensionality is chosen to be $d = 2$ for this application such that a binning can be more easily performed. 5×5 bins are defined in the embedded space as shown in Fig. 7.4a. However, only the ones which contain 5 or more images from each view are selected for the compounding. Next, the 2 second breath-hold of view 2 is rigidly registered to the one of view 1 to obtain a transformation which corrects for the different probe location [Grau et al., 2007]. After adjusting each image from view 2 using this transformation, the compounding algorithm proposed in Yao et al. [2011] is applied to combine the images from each bin.

7.4 Experiments and Results

7.4.1 Experiment 1: Synthetic 4D MR Reconstruction

Synthetic data was generated as described in Appendix B.3 from 6 volunteer scans. In order to evaluate the reconstruction accuracy of the method a volume \tilde{V}_s was reconstructed from each slice s in the dataset, and its \mathcal{L}_2 -distance to the ground-truth volume V_s from which the slice was originally sampled was measured, i.e. $E_s = \|\tilde{V}_s - V_s\|_2$. Since the distribution of E_s was skewed the median reconstruction errors are reported.

In addition to the proposed SAM technique this error was computed for two other image based image reconstruction techniques: simultaneous groupwise manifold alignment (SGA.SIM) as introduced in Chapter 5 and the purely image based method that was also evaluated in the same chapter, and will be referred to here as PIM¹ [Dikaios et al., 2012].

¹PIM was previously called IMBASED.

SGA.REG was not compared to the other methods because of the large computational burden of calculating the registration-based kernel, which would take very long with 100 time points per slice position. Both SGA and PIM compare only neighbouring slice positions and were initially proposed for coronal slices which have smaller changes in appearance between adjacent slice positions. For SGA the optimal parameters proposed in Chapter 5 were used. For SAM the following parameters were chosen by manual tuning: $\mu = 10^{-5}$, $\sigma_1 = 1$, $\sigma_2 = 0.15$, $k = 30$, $d = 4$, $R = 99$. PIM does not have parameters which require tuning.

The reconstruction errors for each of the subjects using the different methods are shown in the box plot in Fig. 7.1a. SAM performed significantly better ($p < 0.001$) than the two other techniques for all volunteers except for volunteer 5 where SGA performed similarly. Significance was assessed using a Wilcoxon signed rank test. SGA and PIM regularly failed to propagate respiratory information across the medial slices. SAM, which embeds all slice positions simultaneously without image comparisons did not suffer from this problem. Coronal views of a single time point of the 4D reconstructions formed from the sagittal input slices for volunteers 2 and 4 are shown in Fig. 7.1b.

7.4.2 Experiment 2: 4D MR Reconstruction from Real Data

Real slice-by-slice data from 4 volunteers were acquired and 4D MR reconstructions were formed as explained in Sec. 7.3.1. Again the method was compared against SGA and PIM using the same parameters as above for all methods.

The first two dimensions of the globally consistent embedded space containing the data from all slice positions is shown in Fig. 7.2. The low-dimensional embeddings corresponding to the image slices acquired at most slice positions are aligned well. Only the embedding corresponding to the left-most slice position is slightly misaligned in the centre region.

On the left-hand side of Fig. 7.3a, examples of coronal slices through single time-frames of the 4D MR reconstruction are shown for volunteer 2. All three methods show reasonable reconstructions. However, as in Experiment 1, the respiratory information did not always propagate through medial slices for SGA and PIM. To illustrate this, lines on sagittal slices from the left and right hemi-diaphragm were manually defined and the 4D MR reconstruction was plotted along these lines for a short time range as shown on the right-hand side of Fig. 7.3a. Because the left and right hemi-diaphragms normally do not move independently (see Chapter 2), in a correct 4D MR reconstruction there should be some correlation between the two curves. For the shown volunteer, SAM has the highest correlation. For the other two techniques respiratory information gets lost in the body middle as before in Experiment 1. As a quantitative measure, curves from such line profiles (see red curves in Fig. 7.3a) were extracted for all volunteers for the entire 4D MR reconstruction and the

NCC between the two curves was calculated for each method. The results are shown in Fig. 7.3b.

7.4.3 Experiment 3: 3D Ultrasound Compounding

US data from two views from one healthy volunteer were acquired and binned as described in Sec. 7.3.2. The following parameter values were used for SAM: $\mu = 10^{-4}$, $\sigma_1 = 1$, $\sigma_2 = 0.2$, $k = 20$, $R = 200$. Two single frames of the two views are shown in Fig. 7.4c. The gating resulted in 8 bins with more than 5 volumes from each view. These are highlighted in green in Fig. 7.4a. In Fig. 7.4b the reconstructions from these 8 bins are shown. In this experiment the SGA and PIM methods were not evaluated because the views are not similar enough in image space to apply those techniques. By observing the liver boundary indicated by the arrow in the first frame, it can be verified that the bins order the data from exhale to inhale. For comparison a compounded volume of all acquired images without gating is shown in Fig. 7.4d in which the same structure appears blurred. To emphasize the 3D nature of the data, in Fig. 7.4e a 3D rendering of bin 1 is shown.

7.5 Discussion

In this chapter, a novel method for self-aligning manifold embeddings was presented which can reduce the dimensionality of multiple datasets into a single globally consistent space. The method is completely unsupervised and requires neither prior correspondences nor comparability of the datasets in high-dimensional space. The technique was applied to two problems from the domain of free breathing 4D image reconstruction: 4D US view compounding and 4D MR reconstruction.

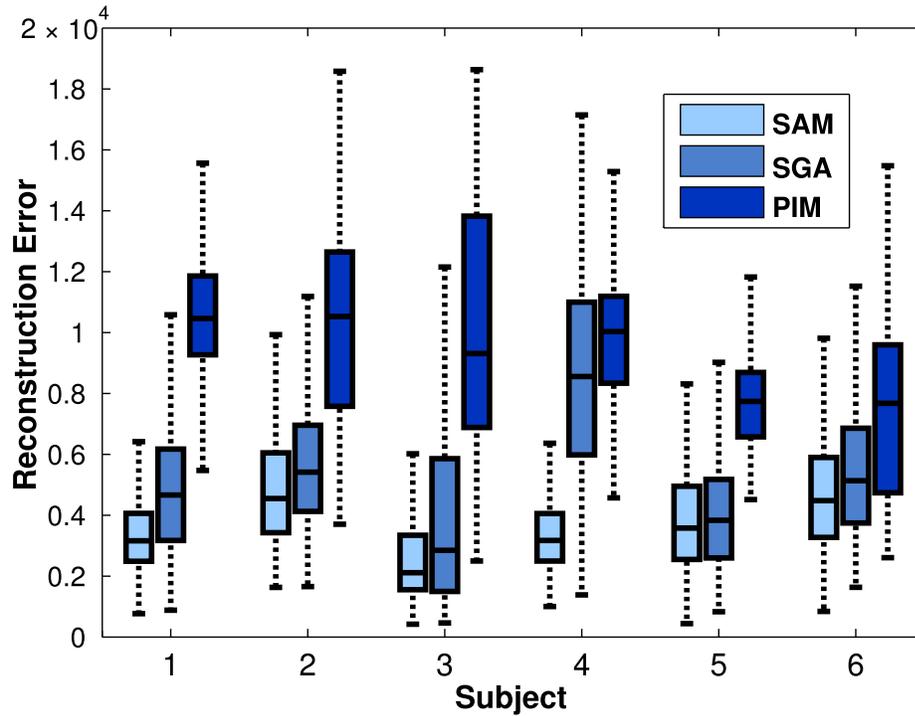
In the 4D MR experiments it was possible to obtain significantly better reconstructions from sagittal slices than using a state-of-the-art image based technique and the technique presented in Chapter 5 on simultaneous groupwise manifold alignment. The 3D US demonstration was restricted to two views for simplicity. However, in principle an arbitrary number of views could be combined in this manner, e.g. to obtain motion images of the entire liver. This would have potential application in radiotherapy planning or motion modelling [McClelland et al., 2013].

One limitation of the proposed technique is that it requires the graph structure to be in some way characteristic of the way the data was generated. This is true for dynamic images acquired during free breathing but it is likely that most real datasets in medical imaging will have such a structure. Nevertheless, one inherent drawback to the method is that it requires enough data points for the graph structure to be representative of the data.

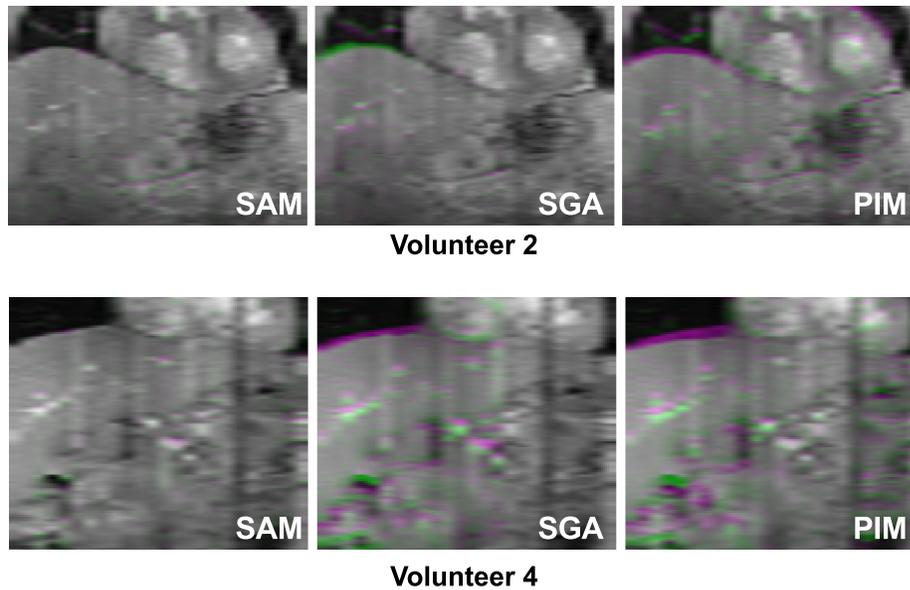
In order to formulate the similarity kernel, work proposed in the graph matching lit-

erature, in particular the work by Gori et al. [2005], was extended. Graph matching and manifold alignment are, in fact, strongly related techniques. The key difference is that in graph matching one is typically interested in a one-to-one correspondence between graph nodes. In MA, on the other hand, a continuous low-dimensional space is obtained in which more complex relations between all the data points can be formed, such the binning which was performed in the US gating experiment.

The idea of self-aligning manifolds may find application in a wide range of problems. Since the method does not require the images to be comparable in high-dimensional space, it could also be applied to matching the motion in multi-modal image pairs, such as PET and MR for motion corrected PET reconstruction [Dikaivos et al., 2012; King et al., 2012], or the alignment of pre-treatment MR or CT images to intra-treatment imaging such as real-time US images for image-guided interventions [McClelland et al., 2013]. Lastly, the potential of this technique is not limited to analysing motion. The globally consistent embedded space of multiple datasets could also be used as input features for further statistical analysis such as classification of Alzheimer’s disease from images acquired with different modalities [Guerrero et al., 2014].



(a) Box plot of reconstruction errors



(b) Example reconstructions

Figure 7.1: Results of synthetic 4D MR reconstruction experiment. (a) Box plot of reconstruction errors. (b) Coronal slices through a single time point of the 4D MR reconstructions obtained using the three methods for two volunteers. The green and purple areas denote positive and negative differences to the ground-truth.

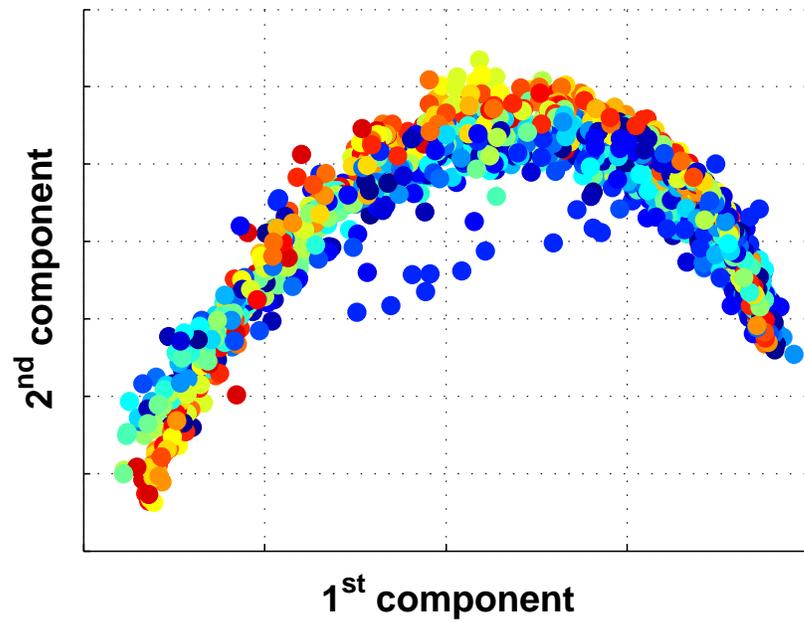
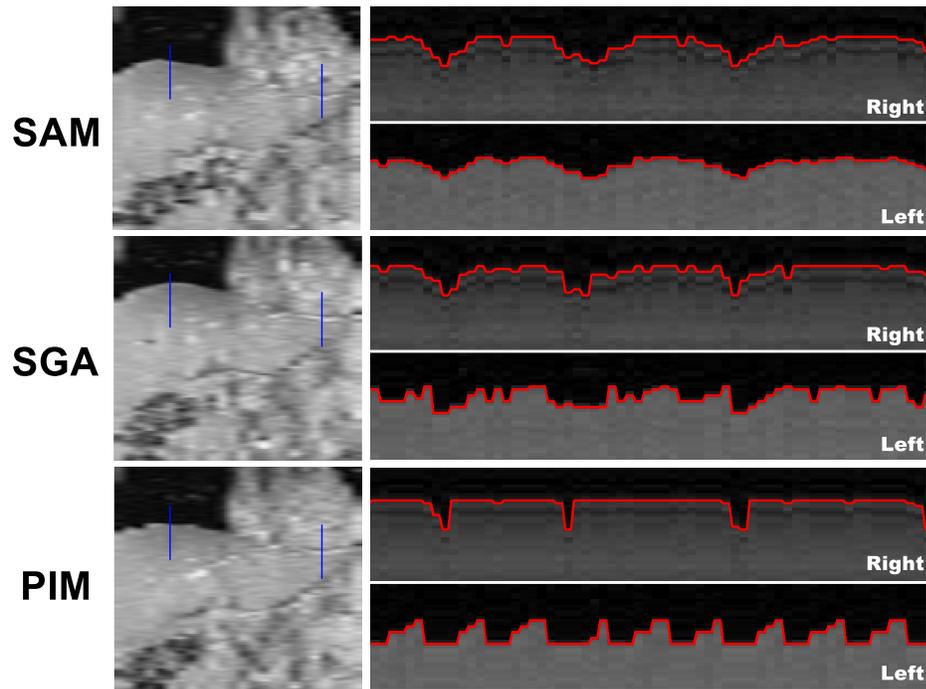
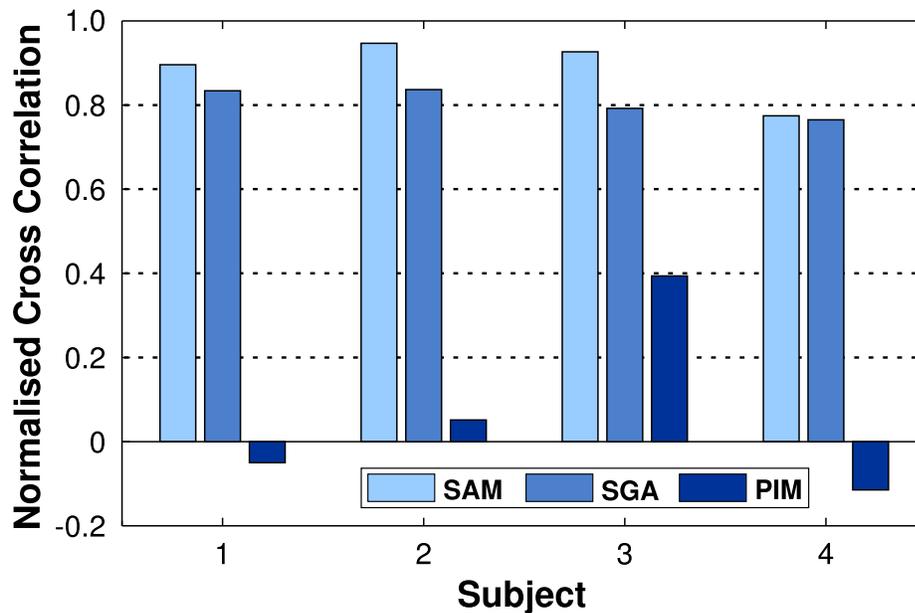


Figure 7.2: The first two components of the aligned manifold embeddings of the data from all MR slice positions, where each colour corresponds to a different slice position. Blue corresponds to the left-most, and red to the right-most slice position.

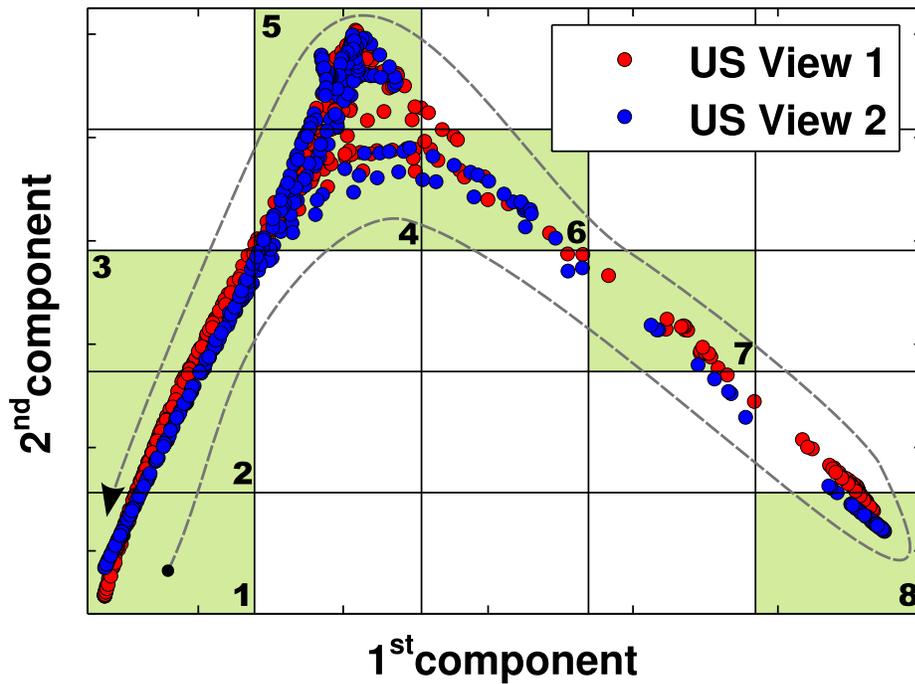


(a) Example reconstructions

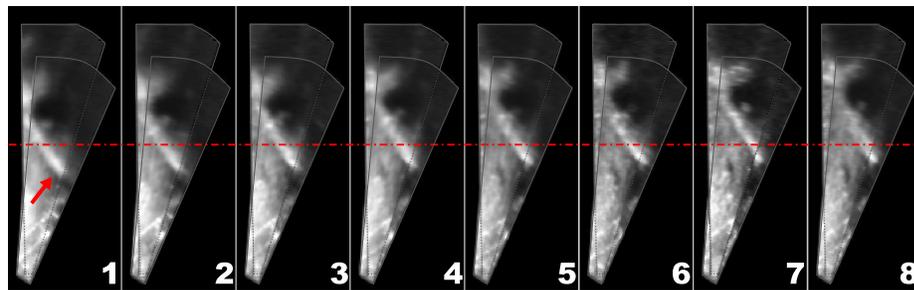


(b) NCC between left and right hemi-diaphragm

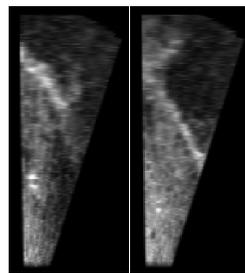
Figure 7.3: Results of the real 4D MR reconstruction. (a) Coronal slice through a single time point of the 4D MR sequence obtained using the three compared methods and line profiles extracted from the left and right hemi-diaphragm from sagittal slices indicated by the blue lines (the right line is at a more posterior position on this slice). (b) NCC between the left and right line profiles over the entire duration of the 4D MR sequence.



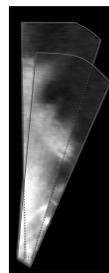
(a)



(b)



(c)



(d)



(e)

Figure 7.4: Results of US experiment. (a) Self-aligned manifolds of the two views, with gating bins. The arrow indicates the path around the manifold during one full breath. (b) Reconstructions using the data in the high-lighted bins, (c) single non-compounded frames from the two views, (d) compounded volume without gating, (e) 3D rendering of compounded bin 1.

Chapter 8

Conclusions

This chapter contains the final discussion and conclusions. A summary of the novel contributions introduced in this thesis is presented in Section 8.1. In Section 8.2, an analysis of the current limitations of the proposed methods is provided, along with possible future directions to address these limitations. Lastly, final remarks are given in Section 8.3.

8.1 Summary of Novel Contributions

Respiratory motion is a significant problem in medical imaging affecting a wide range of modalities, such as MR or PET, and treatments such as image-guided interventions. Manifold learning offers an elegant solution to represent respiratory motion using only a few dimensions. Multiple datasets with limited or no similarity in image space may nevertheless lie on similar manifolds. The nature of these manifolds is defined by how the images in one dataset change with respect to each other due to respiratory motion rather than the absolute appearance of the images. Manifold alignment methods allow correspondences to be established between multiple datasets which lie on similar manifolds. In this thesis manifold alignment approaches were explored to compensate for, or model respiratory motion.

In Chapter 5, manifold alignment was used for the robust data-driven reconstruction of 4D MR volumes from sequentially acquired coronal 2D MR slices. In that chapter, a number of manifold alignment approaches were studied. It was found that a novel groupwise scheme (SGA), where multiple datasets are embedded in groups of two, succeeded in reconstructing 4D MR volumes more accurately than state-of-the-art slice stacking methods, due to its ability to capture inter- and intra-cycle respiratory variabilities, which occur in respiratory motion. Accurate 4D MR volumes have applications in many domains such as the accurate planning of treatments such as radiotherapy to avoid the irradiation of healthy tissue, or motion correction of other simultaneously acquired modalities as was demonstrated in

Chapter 5 for PET/MR. Accurate respiratory motion correction of PET data is gaining importance as the spatial resolution of clinically available PET scanners is improving. This has important implications for the detectability of lesions and correct disease diagnosis and staging.

In Chapter 6, the SGA method was extended to build an autoadaptive motion model which has the ability to adapt to previously unseen breathing patterns during the application phase of MR-guided treatments. To this end SGA was modified to use sagittal 2D MR input slices and data from a single coronal slice position for guidance. It was shown that such a model can successfully improve its motion estimation accuracy over the duration of a typical treatment. Such autoadaptivity could obviate the need for time-consuming model recalibration once a model loses its validity due to changing respiratory patterns. Thus, motion estimation accuracy could be steadily maintained in MR-guided treatments such as MR-guided HIFU, or potentially MR-guided radiotherapy. This may help to spare healthy tissue, and may improve the success rate of such treatments while reducing the time burden for the patients.

Lastly, in Chapter 7, manifold alignment was used to find correspondences between datasets, which unlike the data in Chapters 5 and 6 are not comparable in image space. In particular, the methodology was applied to reconstruct 4D MR volumes from sequentially acquired non-overlapping sagittal 2D MR slices, which differ too much from each other in appearance to be robustly compared in image space, and for gating of 3D ultrasound data from two different views of the same organ. To this end a novel one-step manifold alignment approach was proposed which established similarities between the datasets based on each dataset's internal graph structure. The application to 4D MR reconstruction can be seen as an extension of the method from Chapter 5 to sagittal input slices, which allows the reconstructed volumes to have higher resolution in the plane where most motion occurs. However, the method proposed in this chapter has the potential to be applied to many other datasets which do not share any similarity in image space.

8.2 Current Limitations and Future Directions

The focus of all methods presented in this thesis was on modelling and imaging respiratory motion. However, apart from respiratory motion, there are a number of other sources of motion which impede medical imaging techniques and treatments. For applications such as cardiac imaging or image-guided treatments of the heart, naturally cardiac motion is a significant issue. It is likely that the methods developed here would be applicable directly to sets of images where cardiac motion is the only source of motion. However, in experiments it was found that the techniques are not robust to two independent sources of motion.

In order to focus on the breathing motion for the methods in this thesis either cardiac gated images were used (Chapters 5 and 6) or a region was imaged where cardiac motion is minimal or absent (i.e. the liver in Chapter 7). Extending the manifold alignment techniques developed here to accurately capture multiple sources of motion would be an interesting and important future direction for all of the applications discussed in this thesis: 4D MR volumes could be acquired faster and could potentially also reflect cardiac motion, the MR imaging of PET/MR acquisitions could more easily fit into a typical treatment duration, and an autoadaptive motion model could be updated more frequently and could also potentially be used to model cardiac motion. Some related works have already explored separating cardiac and respiratory motion using manifold learning. In particular, Bhatia et al. [2012b] and Bhatia et al. [2012a] explored a region-wise manifold learning approach for this purpose. Possibly similar methodology could be applied to the techniques developed here.

In this project, mostly T1-weighted 2D MR imaging was used apart from the demonstration on multiple 3D ultrasound views in Chapter 7. Another interesting future direction is the application of the methodology developed here to other imaging modalities. For instance, one possibility for future work may be applying the method to other MR sequences. For example, Andia et al. [2013] proposed a subtractive, blood flow-independent MR technique for visualising the arterial vessel wall. The technique in that paper is based on the acquisition and subtraction of two datasets, one obtained with and one without a T2-preparation prepulse. Currently, it is assured that the two datasets are at the same respiratory state using gating based on a diaphragmatic pencil-beam navigator and one 7 mm gating window. Potentially, the accuracy and scan efficiency of such an approach could be improved by performing gating in a joint embedded space obtained using manifold alignment.

Potential uses of manifold alignment techniques are not restricted to MR imaging. Another application which is conceptually similar to the ones discussed in this thesis is the matching of cardiac and respiratory position in different X-ray fluoroscopy imaging planes for the subsequent 3D localisation of catheters [Panayiotou et al., 2014]. Data from each imaged plane could form a separate dataset as input for the manifold alignment.

In Chapter 7, a manifold alignment method was developed which does not require any resemblance of the datasets in high-dimensional space. Using such an approach potentially even correspondences between data acquired using different modalities could be found, as long as the motion observed is the same in both modalities. In this manner, for example, very distinct data such as PET list mode and MR imaging (or even MR k-space) data could be aligned in a joint low-dimensional space, which could be subsequently used to obtain motion corrected and attenuation corrected PET reconstructions. Another interesting, al-

beit challenging, extension of the present work could be to align intra-operative ultrasound or X-ray fluoroscopy data to pretreatment 4D MR imaging in order to augment the information available to the surgeon. This could be seen as a different type of motion model in which the 4D MR would be the calibration data, and the ultrasound (or fluoroscopy) images would be the surrogate data.

All methods developed in this project were investigated quantitatively and qualitatively using synthetic data, and real volunteer scans. However, none of the works have yet been translated to a clinical setting. The PET/MR motion technique described in Chapter 5 could be directly applied to clinical data in its current form. Indeed, some preliminary PET/MR scans have already been acquired using the proposed MR slice-by-slice acquisition and 4D MR sequences were successfully reconstructed from the data. However, this is still a work in progress.

The autoadaptive motion modelling technique described in Chapter 6 has potential for improvements. For example the MR acquisition protocol could be modified to allow for faster imaging, and efficient parallel implementations of the method could significantly reduce computation times. Both modifications could help to increase the update frequencies at which the model can provide 3D motion estimates. Furthermore, in its current form, the model is not capable of estimating R-L motion. Future work could try to overcome this problem by also incorporating motion estimates from one or multiple coronal slices into the estimated motion fields.

8.3 Final Remarks

This project explored the potential of manifold alignment for overcoming the problem of respiratory motion in medical imaging. A number of related works have explored manifold learning and manifold alignment for analysing motion in medical images. However, most related works either reduced the alignment of manifolds to the simpler case of scaling 1D curves, or applied it only to a small number of datasets. It is much easier to align manifolds in just one dimension, however, a 1D signal is a less powerful descriptor of the data and cannot account for variability in the data such as respiratory inter- and intra-cycle variabilities. Likewise, aligning the embeddings of only two datasets is an easier problem than the general problem of aligning an arbitrary number of embeddings. In this project, a number of powerful novel techniques have been proposed which allow the alignment of multiple datasets in order to establish relations between them in a joint low-dimensional space of adequate dimensionality. It is concluded that manifold alignment methodology has great potential for the data-driven analysis and compensation of respiratory motion, and physiological motion in general, in a wide range of applications. The techniques developed in this project have already been demonstrated for several applications in medical imaging,

but many more applications could be approached using the methodology developed here.

Appendix A

Additional Proofs

A.1 Extension of LLE to Two Datasets

Here, I will show that the augmented LLE cost function for simultaneously embedding two datasets in Eq. (4.6) can be rewritten in matrix form as Eq. (4.9), which is of the same form as the original LLE cost function in matrix form (Eq. (4.3)).

By plugging in all cost functions Eq. (4.6) can be written as

$$\begin{aligned} C_{tot}(Y_p, Y_q) &= \phi(Y_p) + \phi(Y_q) + \mu \cdot \psi(Y_p, Y_q) \\ &= \sum_i |y_p^i - \sum_j W_p^{ij} y_p^j|^2 + \sum_i |y_q^i - \sum_j W_q^{ij} y_q^j|^2 + \mu \cdot \sum_{i,j} |y_p^i - y_q^j|^2 U_{pq}^{ij} \\ &= \sum_{i,j} m_p^{ij} \langle y_p^i, y_p^j \rangle + \sum_{i,j} m_q^{ij} \langle y_q^i, y_q^j \rangle + \\ &\quad \mu \sum_i |y_p^i|^2 \sum_j U_{pq}^{ij} + \mu \sum_j |y_q^j|^2 \sum_i U_{pq}^{ij} - 2\mu \sum_{i,j} U_{pq}^{ij} \langle y_p^i, y_q^j \rangle. \end{aligned}$$

In the last equality above, the LLE cost functions were rewritten in their quadratic form as was shown by Saul and Roweis [2003] and the squared distance in the last term was expanded. Using the following definitions for the sums over the rows and columns of the matrix U_{pq}

$$\begin{aligned} D_{pq}^i &:= \sum_j U_{pq}^{ij} \\ D_{qp}^j &:= \sum_i U_{pq}^{ij} \end{aligned}$$

the cost function can be further simplified to

$$\begin{aligned}
C_{tot}(Y_p, Y_q) &= \sum_{i,j} m_p^{ij} \langle y_p^i, y_p^j \rangle + \sum_{i,j} m_q^{ij} \langle y_q^i, y_q^j \rangle + \\
&\quad \mu \sum_i D_{pq}^i \langle y_p^i, y_p^i \rangle + \mu \sum_j D_{pq}^j \langle y_p^j, y_p^j \rangle - 2\mu \sum_{i,j} U_{pq}^{ij} \langle y_p^i, y_q^j \rangle. \\
&= \text{Tr}(Y_p M_p Y_p^\top) + \text{Tr}(Y_q M_q Y_q^\top) + \\
&\quad \mu \cdot \text{Tr}(Y_p D_{pq} Y_p) + \mu \cdot \text{Tr}(Y_q D_{qp} Y_q^\top) - 2\mu \cdot \text{Tr}(Y_p U_{pq} Y_q^\top) \\
&= \text{Tr}(Y_p M_p Y_p^\top + Y_q M_q Y_q^\top + \mu Y_p D_{pq} Y_p^\top + \mu Y_q D_{qp} Y_q^\top - 2\mu Y_p U_{pq} Y_q^\top)
\end{aligned}$$

In the second to last equality, Lemma A.1.1 was used and the last equality results from the basic properties of the trace operator. D_{pq} and D_{qp} are diagonal matrices containing the values D_{pq}^i and D_{qp}^j , respectively.

Lemma A.1.1. *Given three matrices $A, B \in \mathbb{R}^{d \times N}$ and $H \in \mathbb{R}^{N \times N}$, where \mathbf{a}_i and \mathbf{b}_j are the columns of A and B , respectively, the following statement is true:*

$$\sum_{i,j} H^{ij} \langle \mathbf{a}^i, \mathbf{b}^j \rangle = \text{Tr}(AHB^\top)$$

Proof. Using the definition of the trace operator

$$\begin{aligned}
\text{Tr}(AHB^\top) &= \sum_{k,i,j} A^{ki} H^{ij} (B^\top)^{jk} \\
&= \sum_{i,j} H^{ij} \sum_k A^{ki} B^{kj} \\
&= \sum_{i,j} H^{ij} \langle \mathbf{a}^i, \mathbf{b}^j \rangle \quad \square
\end{aligned}$$

The cost function can be rewritten using the fact that $\text{Tr}(A) = \text{Tr}(A^\top)$ as

$$\begin{aligned}
C_{tot}(Y_p, Y_q) &= \text{Tr}(Y_p M_p Y_p^\top + Y_q M_q Y_q^\top + \mu Y_p D_{pq} Y_p^\top + \mu Y_q D_{qp} Y_q^\top \\
&\quad - \mu Y_p U_{pq} Y_q^\top - \mu Y_q U_{pq}^\top Y_p^\top). \\
&= \text{Tr}(Y_p (M_p + \mu D_{pq}) Y_p^\top + Y_q (M_q + \mu D_{qp}) Y_q^\top - \mu Y_p U_{pq} Y_q^\top - \mu Y_q U_{pq}^\top Y_p^\top).
\end{aligned}$$

From here it can be easily verified that this is equivalent to Eq. (4.9), i.e.,

$$C_{tot}(Y_p, Y_q) = Tr \left(\left[\begin{array}{c} Y_p^\top \\ Y_q^\top \end{array} \right]^\top \left[\begin{array}{cc} M_p + \mu D_{pq} & -\mu U_{pq} \\ -\mu U_{pq}^\top & M_q + \mu D_{qp} \end{array} \right] \left[\begin{array}{c} Y_p^\top \\ Y_q^\top \end{array} \right] \right).$$

Appendix B

Generation of Synthetic Data

In this thesis, various types of synthetic data were used. For Chapter 5, 2D MR coronal slice-by-slice data were simulated by simply transforming a breath-hold volume using motion fields derived from a number of low resolution 3D dynamic volumes. This data is referred to here as TYPE 1 and the details of the data generation are discussed in Section B.1. Furthermore, for the same Chapter synthetic PET/MR data were generated from publicly available 4D CT volumes. The details are discussed in Section B.2.

This type of data, however, had the drawback that each slice position had exactly the same sampling of respiratory positions, which may be unrealistic. To account for this, for Chapters 6 and 7, a more realistic type of synthetic 2D MR slice-by-slice data (referred to here as TYPE 2) was used. Given the requirements of the respective experiments this data had sagittal slice orientation, with one additional coronal slice position for the experiments in Chapter 6. The generation involved first creating a simple motion model to generate a number of dense 3D motion fields. From this, synthetic 2D MR image data was generated for Chapter 6 as described in Section B.3.1. To generate the synthetic 2D motion fields used in Chapter 6 the synthetic slice data from each slice position were registered to a reference exhale volume, which is described in Section B.3.2.

Note that, although the TYPE 2 data is more realistic than TYPE 1, it is significantly more computationally expensive to generate, with processing times of multiple days for one volunteer.

B.1 Generation of Simple Synthetic MR Data (TYPE 1)

In order to generate the synthetic data of TYPE 1, a slice-by-slice breath-hold volume consisting of coronal slices was acquired at end-exhale covering all slice positions using the same slice-by-slice acquisition sequence that was used for the real data described in Section 5.2.1. In addition, for each volunteer, 50 low-resolution dynamic volumes were acquired on

the same Philips Achieva 3T MR system using a cardiac-triggered T1-weighted gradient echo sequence with an acquired image resolution of $1.5 \times 4.1 \times 5 \text{ mm}^3$, and an acquisition time of approximately 600 ms. The highest resolution was chosen in the S-I direction, where most respiratory motion occurs [Seppenwoolde et al., 2002]. An example of a coronal slice from such a volume was shown in Figure 3.1a. In order to compare the developed methods in Chapter 5 against the PBNV technique, a 1D pencil beam navigator was recorded immediately before and after acquiring each dynamic image. To account for the length of the 3D acquisition the leading and trailing pencil beam navigators were averaged to arrive at a better estimate of respiratory position. Note that for this type of synthetic data the pencil beam navigator was used only to implement the PBNV technique proposed by Würslin et al. [2013]. In the next step, a reference exhale image was chosen manually from the 50 low-resolution dynamic images. Next, motion fields were obtained for each of the dynamic images by registering them to the reference volume using B-spline registration. Finally, the breath-hold slices were transformed using the motion fields to arrive at synthetic slices at different respiratory positions. However, note that while the deformations of the thorax overall are realistic, the motion estimates inside the lungs are not reliable as they were derived from low-resolution volumes with little contrast in this area.

B.2 Generation of Synthetic PET Data

Based on a breath-hold MR slice-by-slice scan of volunteer A in Chapter 5 (a slice of the volume is shown in Figure B.1a), a 3D FDG uptake map with realistic SUVs was manually created (Figure B.1c). For the purpose of evaluating the accuracy of motion correction inside the lungs, additionally an artificial lung tumour with 1.3 cm diameter and an SUV of 7 was added to the FDG maps. To simulate the deformations due to respiratory motion these maps were transformed using a set of *generating* motion fields¹. To obtain realistic motion estimates in the whole thorax, including the lung, the generating motion fields were derived from patient 2 of the publicly available POPI 4D CT dataset [Vandemeulebroucke et al., 2011] using B-spline registration and were transported to the coordinate system of our volunteer using the method described in Rao et al. [2002]. A coronal slice of the CT volume transformed to the MR coordinate frame is shown in Figure B.1b. Because the 10 available motion states in the POPI dataset were not enough for our simulation the motion fields were linearly interpolated to arrive at 30 (15 exhaling, 15 inhaling) motion states. The CT volume corresponding to end-exhale was then transformed by the 30 generating motion fields. These CT volumes served as attenuation maps in the following. Based on the 30 motion fields, 30 gates of synthetic PET data ($2 \times 2 \times 2 \text{ mm}^3$ voxels, and 1.67

¹Note that those generating motion fields point in the opposite direction of the *backwards* motion fields estimated in the experiment in Section 5.3.3.

million coincidence events per gate) were simulated using the STIR package [Thielemans et al., 2006] with the ordered subsets expectation maximisation (OSEM) reconstruction algorithm (23 subsets, 11 iterations). These gates were simulated using the 3D FDG uptake map transformed using the generating motion fields. In addition, the transformed CT attenuation maps were used to introduce attenuation effects into the PET simulations and also to correct for such effects in the reconstructions. A single gate of synthetic PET data at end-exhale is shown in Figure B.1d. The tumour is indicated with an arrow. In addition, also 30 PET gates without motion corruption were generated which served as the ground-truth for the evaluation. Lastly, 30 synthetic slice-by-slice MR volumes were obtained by transforming the breath-hold slice-by-slice volume using the generating motion fields. This is the equivalent of simultaneously acquired slice-by-slice data in a real PET/MR scenario. Note that while the synthetic PET and MR data contains intra-cycle variabilities because inhaling and exhaling states were separated in the generating 4D CT data, it does not contain inter-cycle variabilities because in the 4D CT data each time point is an average of multiple breathing cycles.

B.3 Generation of Motion Model-Based Synthetic MR Data (TYPE 2)

In Chapter 5 synthetic data was generated by deriving motion fields from low-resolution 3D MR volumes and then transforming a breath-hold scan using these motion fields as was described above in Section B.1. The drawback of this approach is that every slice position has exactly the same sampling of motion states, which is unrealistic and may artificially oversimplify the manifold alignment problem. In the following two sections an alternative method for generating synthetic data is presented. Note that in order to match the real acquisitions in the respective chapters, in this framework synthetic non-overlapping sagittal MR slices were generated as opposed to the overlapping coronal slices generated for the TYPE 1 data. For the experiments in Chapter 6, additionally data from a single coronal slice position was generated.

The underlying idea of the TYPE 2 synthetic data generation framework was to first build a simple linear subject specific motion model based on two 1D navigators and 3D motion fields derived from a short dynamic low-resolution 3D MR scan containing the type of breathing motion that one wants to synthesise. In Chapter 7, only normal free breathing data was synthesised. However, in Chapter 6, additionally a synthetic dataset based on a deep breathing scan was generated in order to investigate changes in breathing pattern. Note that the motion model used to generate the synthetic data is completely unrelated to the autoadaptive motion model proposed in Chapter 6. By generating random samples

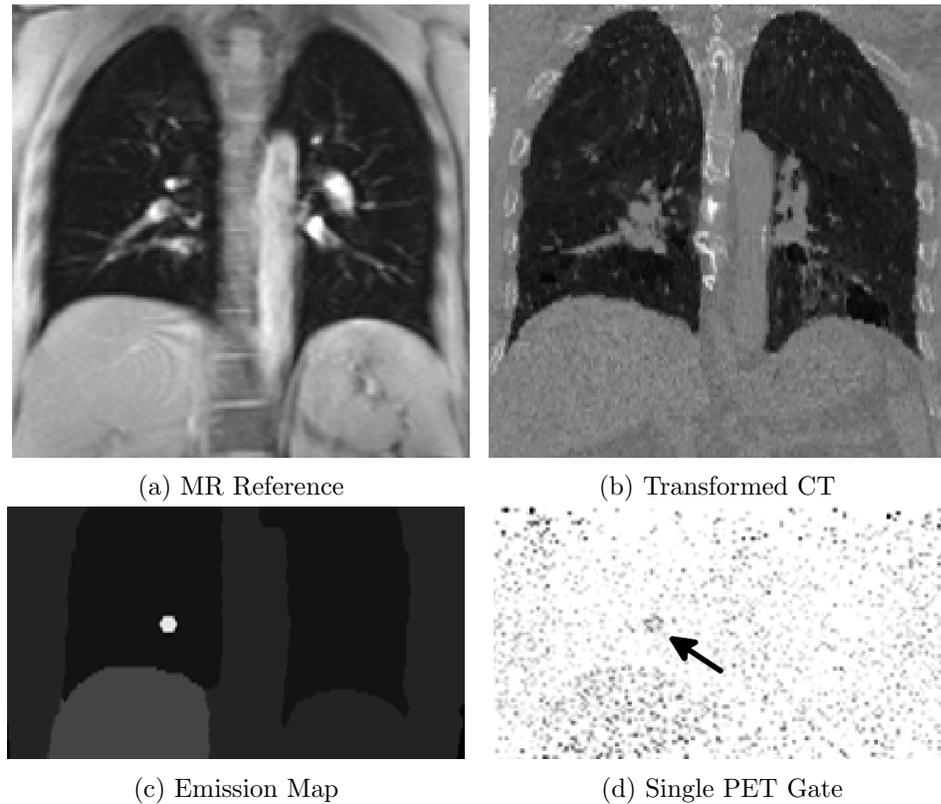


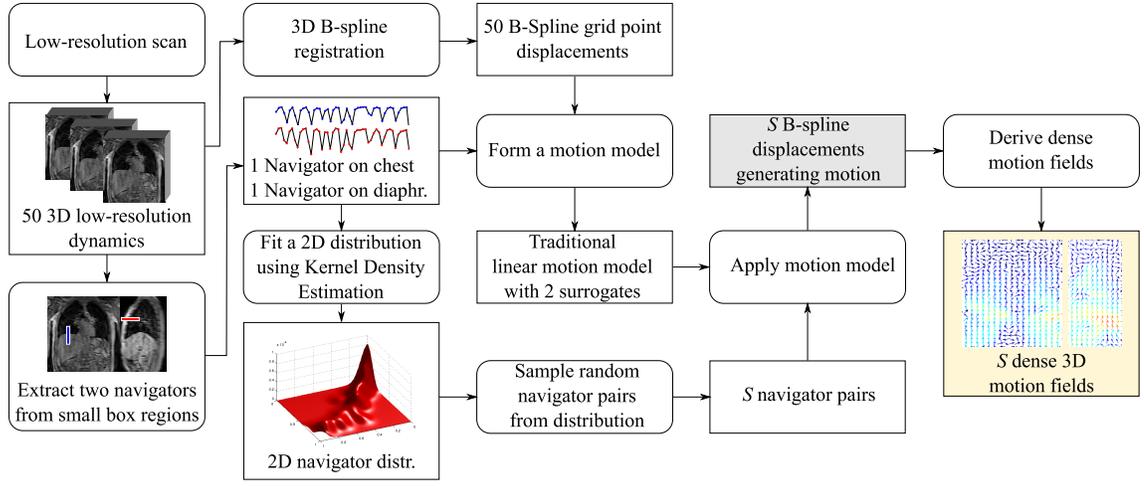
Figure B.1: Coronal slices at the same A-P position through components of the synthetic PET simulation: (a) Reference MR scan for volunteer A, (b) exhale gate from POPI dataset warped to the coordinate system of the volunteer, (c) manually drawn emission map with an artificial tumour inserted in the lower lung (shown in white), and (d) an example of a single synthetic PET gate. The tumour is highlighted with an arrow.

of synthetic navigator values and using them as input to the motion model synthetic, but realistic, respiratory motion deformations could be obtained. These, on one hand, served as a ground-truth for the experiments on autoadaptive motion modelling in Chapter 6, and on the other hand, were used to generate synthetic slice-by-slice data by transforming a slice-by-slice breath-hold scan. This approach had two main advantages: 1) more realistic sampling of respiratory positions, 2) the ability to generate datasets of arbitrary size from a 50 second scan. However, the approach was very computationally expensive.

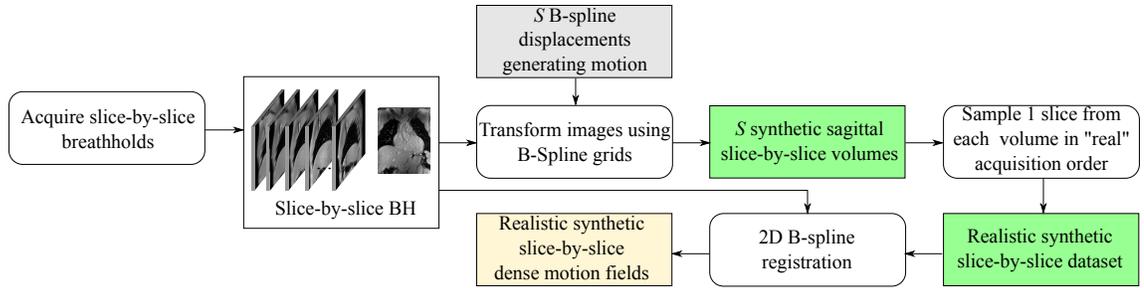
In the following each of the steps outlined above is explained in detail. The generation of the synthetic data is summarised in Fig. B.2. The description of the generation is divided into two parts:

1. The generation of realistic ground-truth motion (see Fig. B.2a).
2. The generation of synthetic slice-by-slice images and the derivation of slice-by-slice

motion fields from them (see Fig. B.2b).



(a) Derivation of ground-truth motion fields



(b) Derivation of synthetic slice-by-slice data

Figure B.2: Overview of synthetic slice-by-slice data generation of TYPE 2. (a) Generation of ground-truth motion fields from a 3D low-resolution MR scan, (b) generation of synthetic slice-by-slice data by applying the ground-truth motion to slice data acquired at end-exhale, and derivation of motion fields. The objects highlighted in green represent the synthetic data and ground-truth for Chapter 6, and the objects highlighted in yellow the synthetic data and ground-truth used for the autoadaptive motion model in Chapter 6. The B-spline displacements highlighted in grey are the connection between Figures (a) and (b) and are the same in both.

For the generation of the ground-truth data, in a first step 50 3D low-resolution MR dynamic images were acquired using the same sequence as for the TYPE 1 data described in Section B.1. These 3D MR images were acquired either under free breathing or under deep breathing. From these volumes 50 control point grid displacements were derived using B-spline registration, with 3 hierarchy levels, a final grid spacing of 15 mm in each direction and no bending energy penalty term. Furthermore, two series of 50 navigator signals s_1, s_2 were extracted from the images by measuring the displacements of small rectangular regions on the dome of the left hemidiaphragm and the anterior chest wall [Savill et al.,

2011]. Two signals were chosen to increase the amount of respiratory variabilities captured in the resulting model. Next, a motion model was formed by fitting a linear function of the navigator signal values to the displacements of each B-spline grid point [McClelland et al., 2013], i.e.

$$\mathbf{v}(\mathbf{t}) = \alpha_1(\mathbf{t}) + \alpha_2(\mathbf{t})s_1 + \alpha_3(\mathbf{t})s_2, \quad (\text{B.1})$$

where $\alpha_1, \alpha_2, \alpha_3$ are the parameters of the motion model and $\mathbf{v}(\mathbf{t})$ are the grid displacements at grid location \mathbf{t} .

In the next step, a 2D distribution was fitted to the navigator signal values using kernel density estimation [Rosenblatt et al., 1956]. Then S random navigator value pairs \tilde{s}_1, \tilde{s}_2 were sampled from this distribution. In the final synthetic data each 2D slice was associated with a ground-truth 3D motion field. Hence, as many synthetic navigator values were needed as the total number of synthetic 2D slices desired in the dataset. The number S was different for the data used in Chapter 6 and 7, and will be discussed below. Next, by substituting the sampled values \tilde{s}_1, \tilde{s}_2 into Eq. (B.1) S synthetic B-spline transformations were obtained, which were then used for the generation of the ground-truth 3D motion fields as well as for the generation of the synthetic slice-by-slice data. The ground-truth motion fields were derived simply by interpolating a dense motion field using the voxel sizes of the slice-by-slice breath-hold volumes described below.

In the remainder of this Appendix, the synthetic slice-by-slice data generation from the synthetic B-spline displacements will be discussed (see Fig. B.2b). In addition to the low-resolution volumes, all sagittal slice positions and one coronal slice position were acquired in two exhale breath-hold acquisitions using the same slice-by-slice protocol as for the real data, which was described in Section 6.2.1. The breath-hold data was then transformed using the synthetic B-spline grid displacements. This led to a sequence of S synthetic slice-by-slice *volumes*. From these volumes then single slices were sampled to create the final synthetic slice-by-slice data. At this point the exact approach differed for the synthetic 2D MR image data used in Chapter 6, and the synthetic motion fields used in Chapter 6. Both will be discussed separately below.

B.3.1 Synthetic 2D MR Slices

For the evaluations in Chapter 6, only the sagittal slice-by-slice data was required. $S = 200$ random navigator pairs, \tilde{s}_1, \tilde{s}_2 , were sampled from the joint navigator value distribution, to obtain 200 synthetic B-spline transformations. These transformations were then applied to the slice-by-slice breath-hold volume to obtain 200 synthetic slice-by-slice volumes. Lastly, for each slice position $\tau_p = 100$ different time points were sampled from each slice position without replacement, generating the final synthetic dataset. The number of slice positions L of each dataset depended on the volunteer scan but was typically around 30. Note that

it was recorded from which of the 200 volumes each slice was sampled such that each slice was associated with a ground-truth reconstruction. The synthetic slice-by-slice data and the ground-truth volumes are highlighted in green in Figure B.2.

B.3.2 Synthetic Motion Fields

For the motion modelling experiments in Chapter 6 a larger dataset was chosen, where no respiratory state occurs twice in the whole dataset. The aim was to generate τ_p independent slices per slice position, which means $S = L \cdot \tau_p$ navigator values needed to be sampled from the navigator value distribution, where L is the total number of slice positions in the synthetic sequence. In the experiments typically L was around 30, and $\tau_p = 50$. In the real acquisitions at each time point only one slice position can be observed, and hence, here only one slice was sampled from each of these volumes. This sampled data constituted the synthetic slice-by-slice dataset and was the synthetic equivalent to the data obtained from a real slice-by-slice scan described in Section 6.2.1.

As part of the model calibration phase in Chapter 6, the slice-by-slice image data was then registered in 2D to the corresponding breath-hold slices in order to obtain slice-by-slice 2D motion fields which are the input to the proposed AAMM. The parameters used for the registration were the same as in Section 6.2.1. Note that each of the sampled slices is associated with exactly one of the generating motion fields. The slice-by-slice motion fields and the motion ground-truth (i.e. the generating motion fields) are highlighted in yellow in Figure B.2a.

Appendix C

Summary of data used in this paper

In this Section an overview of all datasets used in this thesis is given. All 2D MR slice-by-slice data was acquired on a Philips Achieva 3T system using the following parameters: T1-weighted GRE sequence with an acquired in-plane image resolution of $1.4 \times 1.4 \text{ mm}^2$, a slice thickness of 8 mm, repetition and echo times (TR and TE) of 3.1 and 1.9 ms, a flip angle (FA) of 30 deg, and a SENSE factor of 2.

A number of attributes of this sequence were adjusted for the respective experimental requirements of each chapter. In particular, the varying attributes were: slice orientation, slice overlap, cardiac gating, number of slice positions and number of slices per slice position. Table C.1 summarises those attributes for all of the experiments using slice-by-slice MR data performed in this thesis, along with the same attributes for the synthetic data and, as far as they are applicable, the ultrasound data used in Chapter 7.

	Type	# Volunteers	Synth./Real	Orient.	Overlap	Cardiac Gated	L	τ_p
Chapter 5								
Experiments 1 & 3	MR	10	Synth. (TYPE 1)	Coronal	4 mm	Yes	~ 32	50
Experiments 2 & 4	MR	10	Real	Coronal	4 mm	n/a	~ 32	50
Experiment 5	PET/MR	1	Synth.	Coronal	4 mm	n/a	29	30
Chapter 6								
Experiment 1	MR	6	Synth. (TYPE 2)	Sag. & Cor.	0 mm	n/a	~ 30	50
Experiment 2	MR	4	Real	Sag. & Cor.	0 mm	Yes	~ 30	40
Chapter 7								
Experiment 1	MR	6	Synth. (TYPE 1)	Sagittal	0 mm	n/a	~ 30	100
Experiment 2	MR	4	Real	Sagittal	0 mm	No	~ 30	100
Experiment 3	US	1	Real	n/a	n/a	No	2	300

Table C.1: Summary of all the data used in this thesis and their changing attributes organised by chapter. L stands for the number of slice positions, or views, acquired. τ_p denotes how many times each slice position, or view, was sampled.

Bibliography

- N.A. Ablitt, J. Gao, J. Keegan, L. Stegger, D.N. Firmin, and G.Z. Yang. Predictive cardiac motion modeling and correction with partial least squares regression. *IEEE T Med Imaging.*, 23(10):1315–1324, 2004.
- A.P. Aitken, M. Henningsson, R.M. Botnar, T. Schaeffter, and C. Prieto. 100% efficient three-dimensional coronary MR angiography with two-dimensional beat-to-beat translational and bin-to-bin affine motion correction. *Magn Reson Med*, 2014.
- M.E. Andia, M. Henningsson, T. Hussain, A. Phinikaridou, A. Protti, G. Greil, and R.M. Botnar. Flow-independent 3D whole-heart vessel wall imaging using an interleaved T2-preparation acquisition. *Magn Reson Imag.*, 69(1):150–157, 2013.
- P. Arnold, F. Preiswerk, B. Fasel, R. Salomir, K. Scheffler, and P.C. Cattin. 3D organ motion prediction for MR-guided high intensity focused ultrasound. In *Proc. MICCAI*, pages 623–630. Springer, 2011.
- V. Auboiroux, L. Petrusca, M. Viallon, A. Muller, S. Terraz, R. Breguet, X. Montet, C.D. Becker, and R. Salomir. Respiratory-gated MRgHIFU in upper abdomen using an MR-compatible in-bore digital camera. *Biomed Res Int*, 2014, 2014.
- W. Bai and M. Brady. Regularized B-spline deformable registration for respiratory motion correction in PET images. *Phys Med Biol.*, 54(9):2719, 2009.
- C.F. Baumgartner, C. Kolbitsch, J.R. McClelland, D. Rueckert, and A.P. King. Groupwise simultaneous manifold alignment for high-resolution dynamic MR imaging of respiratory motion. In *Proc. IPMI*, pages 232–243. Springer, 2013.
- C.F. Baumgartner, C. Kolbitsch, D.R. Balfour, P.K. Marsden, J.R. McClelland, D. Rueckert, and A.P. King. High-resolution dynamic MR imaging of the thorax for respiratory motion correction of PET using groupwise manifold alignment. *Med Image Anal.*, 18(7): 939–952, 2014a.
- C.F. Baumgartner, C. Kolbitsch, J.R. McClelland, D. Rueckert, and A.P. King. Autoadaptive motion modelling. In *Proc. IEEE ISBI*, pages 457–460. IEEE, 2014b.

- C.F. Baumgartner, A. Gomez, L.M. Koch, J. Housden, C. Kolbitsch, J.R. McClelland, D. Rueckert, and A.P. King. Self-aligning manifolds for matching disparate medical image datasets. In *Proc. IPMI*, page (in press). Springer, 2015.
- M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.*, 15(6):1373–1396, 2003.
- T. Beyer, D.W. Townsend, T. Brun, P.E. Kinahan, M. Charron, R. Roddy, J. Jerin, J. Young, L. Byars, R. Nutt, et al. A combined PET/CT scanner for clinical oncology. *J Nucl Med.*, 41(8):1369–1379, 2000.
- K.K. Bhatia, A.N. Price, J.V. Hajnal, and D. Rueckert. Localised manifold learning for cardiac image analysis. In *Proc. SPIE.*, pages 83140H–83140H–9, 2012a.
- K.K. Bhatia, A. Rao, A.N. Price, R. Wolz, J. Hajnal, and D. Rueckert. Hierarchical manifold learning. In *Proc. MICCAI.*, pages 512–519. Springer, 2012b.
- J.M. Blackall, S. Ahmad, M.E. Miquel, J.R. McClelland, D.B. Landau, and D.J. Hawkes. MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. *Phys Med Biol.*, 51:4147–4169, 2006.
- R.A. Bundschuh, A. Martínez-Moeller, M. Essler, M-J. Martínez, S.G. Nekolla, S.I. Ziegler, and M. Schwaiger. Postacquisition detection of tumor motion in the lung and upper abdomen using list-mode PET data: a feasibility study. *J Nucl Med*, 48(5):758–763, 2007.
- J.T. Bushberg, J.A. Seibert, E.M. Leidholdt, and J.M. Boone. *The essential physics of medical imaging*. Lippincott Williams & Wilkins, 2001.
- F. Camastra and A. Vinciarelli. Estimating the intrinsic dimension of data with a fractal-based method. *IEEE T Pattern Anal.*, 24(10):1404–1407, 2002.
- R. Colgan, J. McClelland, D. McQuaid, P.M. Evans, D. Hawkes, J. Brock, D. Landau, and S. Webb. Planning lung radiotherapy using 4D CT data and a motion model. *Phys Med Biol*, 53(20):5815, 2008.
- G. Cruz, D. Atkinson, C. Buerger, T. Schaeffter, and C. Prieto. Accelerated motion corrected three-dimensional abdominal MRI using total variation regularized sense reconstruction. *Magn Reson Med*, 2015.
- P.G. Danias, M.V. McConnell, V.C. Khasgiwala, M.L. Chuang, R.R. Edelman, and W.J. Manning. Prospective navigator correction of image position for coronary MR angiography. *Radiology*, 203(3):733–736, 1997.

- M. Dawood, N. Lang, X. Jiang, and K.P. Schafers. Lung motion correction on respiratory gated 3-D PET/CT images. *IEEE T Med Imaging.*, 25(4):476–485, 2006.
- L. Dietrich, S. Jetter, T. Tücking, S. Nill, and U. Oelfke. Linac-integrated 4D cone beam CT: first experimental results. *Phys Med Biol*, 51(11):2939, 2006.
- N. Dikaios, D. Izquierdo-Garcia, M.J. Graves, V. Mani, Z.A. Fayad, and T.D. Fryer. MRI-based motion correction of thoracic PET: initial comparison of acquisition protocols and correction strategies suitable for simultaneous PET/MRI systems. *Eur radiol.*, 22(2):439–446, 2012.
- K. Ding, Y. Yin, K. Cao, G.E. Christensen, C-L. Lin, E.A. Hoffman, and J.M. Reinhardt. Evaluation of lobar biomechanics during respiration using image registration. In *Proc. MICCAI*, pages 739–746. Springer, 2009.
- D.L. Donoho and C. Grimes. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *P Natl Acad Sci USA.*, 100(10):5591–5596, 2003.
- R.L. Ehman and J.P. Felmlee. Adaptive technique for high-definition MR imaging of moving structures. *Radiology*, 173(1):255–263, 1989.
- R.L. Ehman, M.T. McNamara, M. Pallack, H. Hricak, and C.B. Higgins. Magnetic resonance imaging with respiratory gating: techniques and advantages. *Am J Roentgenol*, 143(6):1175–1182, 1984.
- J. Ehrhardt, R. Werner, A.S. Richberg, and H. Handels. Statistical modeling of 4D respiratory lung motion using diffeomorphic image registration. *IEEE T Med Imaging*, 30(2):251–265, 2011.
- A. Faller and M. Schünke. *The human body: an introduction to structure and function*. Thieme, 2004.
- H. Fayad, T. Pan, C. Roux, and D. Visvikis. A generic respiratory motion model for motion correction in PET/CT. In *IEEE Nucl Sci Med Image Conf*, pages 2455–2458. IEEE, 2010.
- P. Fischer, T. Pohl, and J. Hornegger. Real-time respiratory signal extractino from X-ray sequences using incremental manifold learning. In *Proc. IEEE ISBI.*, pages 915–918. IEEE, 2014.
- G. Gao, J. McClelland, S. Tarte, J.M. Blackall, and D.J. Hawkes. Modelling the respiratory motion of the internal organs by using canonical correlation analysis and dynamic MRI. *Workshop on Pulmonary Image Analysis at MICCAI*, 2008.

- M. Georg, R. Souvenir, A. Hope, and R. Pless. Manifold learning for 4D CT reconstruction of the lung. In *Proc. IEEE CVPRW.*, pages 1–8, 2008.
- M. Gori, M. Maggini, and L. Sarti. Exact and approximate graph matching using random walks. *IEEE T Pattern Anal.*, 27(7):1100–1111, 2005.
- V. Grau, H. Becher, and J.A. Noble. Registration of multiview real-time 3-D echocardiographic sequences. *IEEE T Med Imaging.*, 26(9):1154–65, 2007.
- R. Grimm, S. Fürst, I. Dregely, C. Forman, J.M. Hutter, S.I. Ziegler, S. Nekolla, B. Kiefer, M. Schwaiger, J. Hornegger, et al. Self-gated radial mri for respiratory motion compensation on hybrid pet/mr systems. In *Proc. MICCAI*, pages 17–24. Springer, 2013.
- R. Guerrero, C. Ledig, and D. Rueckert. Manifold alignment and transfer learning for classification of Alzheimer’s disease. In *Proc. MLMI.*, pages 77–84. 2014.
- T. Guerrero, K. Sanders, E. Castillo, Y. Zhang, L. Bidaut, T. Pan, and R. Komaki. Dynamic ventilation imaging from four-dimensional computed tomography. *Phys Med Biol.*, 51(4):777, 2006.
- J. Ham, D. Lee, and L. Saul. Semisupervised alignment of manifolds. In *AI and Statistics*, volume 10, pages 120–127, 2005.
- M.S. Hansen, D. Atkinson, and T.S. Sorensen. Cartesian SENSE and k-t SENSE reconstruction using commodity graphics hardware. *Magn Reson Med.*, 59(3):463–468, 2008.
- D.J. Hawkes, D. Barratt, J.M. Blackall, C. Chan, P.J. Edwards, K. Rhode, G.P. Penney, J. McClelland, and D.L.G. Hill. Tissue deformation and shape models in image-guided interventions: a discussion paper. *Med Im Anal*, 9(2):163–175, 2005.
- M. Hoogeman, J-B. Prévost, J. Nuyttens, J. Pöll, P. Levendag, and B. Heijmen. Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files. *Int J Radiat Oncol.*, 74(1):297–303, 2009.
- K. Hynynen, W.R. Freund, H.E. Cline, A.H. Chung, R.D. Watkins, J.P. Vetro, and F.A. Jolesz. A clinical, noninvasive, MR imaging-monitored ultrasound surgery method. *Radiographics*, 16(1):185–195, 1996.
- M. Isaksson, J. Jalden, and M.J. Murphy. On using an adaptive neural network to predict lung tumor motion during respiration for radiotherapy applications. *Med Phys.*, 32(12):3801–3809, 2005.

- M. Katayama, T. Masui, S. Kobayashi, T. Ito, M. Takahashi, H. Sakahara, A. Nozaki, and H. Kabasawa. Fat-suppressed T2-weighted MRI of the liver: Comparison of respiratory-triggered fast spin-echo, breath-hold single-shot fast spin-echo, and breath-hold fast-recovery fast spin-echo sequences. *J Magn Reson Imaging*, 14(4):439–449, 2001.
- P.J. Keall, G.S. Mageras, J.M. Balter, R.S. Emery, K.M. Forster, S.B. Jiang, J.M. Kapatoes, D.A. Low, M.J. Murphy, B.R. Murray, et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76a). *Med Phys.*, 33(10):3874–3900, 2006.
- A.P. King, R. Boubertakh, K.S. Rhode, Y.L. Ma, P. Chinchapatnam, G. Gao, T. Tangcharoen, M. Ginks, M. Cooklin, J.S. Gill, et al. A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures. *Med Image Anal.*, 13(3):419–431, 2009a.
- A.P. King, K.S. Rhode, R.S. Razavi, and T.R. Schaeffter. An adaptive and predictive respiratory motion model for image-guided interventions: Theory and first clinical application. *IEEE T Med Imaging.*, 28(12):2020–2032, 2009b.
- A.P. King, C. Tsoumpas, C. Buerger, V. Schulz, P. Marsden, and T. Schaeffter. Real-time respiratory motion correction for simultaneous PET-MR using an MR-derived motion model. In *IEEE Nucl Sci Med Image Conf*, pages 3589–3594. IEEE, 2011.
- A.P. King, C. Buerger, C. Tsoumpas, P.K. Marsden, and T. Schaeffter. Thoracic respiratory motion estimation from MRI using a statistical model and a 2-D image navigator. *Med Image Anal.*, 16(1):252–264, 2012.
- V.R. Kini, S.S. Vedam, P.J. Keall, S. Patil, C. Chen, and R. Mohan. Patient training in respiratory-gated radiotherapy. *Med Dosim*, 28(1):7–11, 2003.
- H.U. Klemm, D. Steven, C. Johnsen, R. Ventura, T. Rostock, B. Lutomsky, T. Risius, T. Meinertz, and S. Willems. Catheter motion during atrial ablation due to the beating heart and respiration: Impact on accuracy and spatial referencing in three-dimensional mapping. *Heart Rhythm*, 4(5):587–592, 2007.
- M.O. Köhler, B. Denis de Senneville, B. Quesson, C.T.W. Moonen, and M. Ries. Spectrally selective pencil-beam navigator for motion compensation of MR-guided high-intensity focused ultrasound therapy of abdominal organs. *Magn Reson Med.*, 66(1):102–111, 2011.
- H.W. Korin, R.L. Ehman, S.J. Riederer, J.P. Felmlee, and R.C. Grimm. Respiratory kinematics of the upper abdominal organs: a quantitative study. *Magn Reson Med*, 23(1):172–178, 1992.

- D-J. Kroon. *Segmentation of the mandibular canal in cone-beam CT data*. PhD thesis, University of Twente, Enschede, The Netherlands, 2011.
- A. Kubias, F. Deinzer, T. Feldmann, D. Paulus, B. Schreiber, and T. Brunner. 2D/3D image registration on the GPU. *S Mach Perc*, 18(3):381–389, 2008.
- H.W. Kuhn. The Hungarian method for the assignment problem. *Nav Res Logist Q.*, 2(1-2):83–97, 1955.
- S. Lafon and A.B. Lee. Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization. *IEEE T Pattern Anal.*, 28(9):1393–1403, 2006.
- F. Lamare, M.J.L. Carbayo, T. Cresson, G. Kontaxakis, A. Santos, C.C. Le Rest, A.J. Reader, and D. Visvikis. List-mode-based reconstruction for respiratory motion correction in PET using non-rigid body transformations. *Phys Med Biol*, 52(17):5187, 2007.
- G. Langs, D. Lashkari, A. Sweet, Y. Tie, L. Rigolo, A.J. Golby, and P. Golland. Learning an atlas of a cognitive process in its functional geometry. In *Proc. IPMI*, pages 135–146. Springer, 2011.
- G. Langs, A. Sweet, D. Lashkari, Y. Tie, L. Rigolo, A.J. Golby, and P. Golland. Decoupling function and anatomy in atlases of functional connectivity patterns: Language mapping in tumor patients. *NeuroImage*, 103:462–475, 2014.
- M. Lewandowski, J. Martinez-del Rincon, D. Makris, and J.C. Nebel. Temporal extension of Laplacian eigenmaps for unsupervised dimensionality reduction of time series. In *Proc. ICPR.*, 2010.
- A.S. Lichter and T.S. Lawrence. Recent advances in radiation oncology. *New Engl J Med*, 332(6):371–379, 1995.
- C. Liu, Larry A. Pierce I.I., A.M. Alessio, and P.E. Kinahan. The impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging. *Phys Med Biol*, 54(24):7345, 2009.
- Y.L. Liu, S.J. Riederer, P.J. Rossman, R.C. Grim, J.P. Debbins, and R.L. Ehman. A monitoring, feedback, and triggering system for reproducible breath-hold MR imaging. *Magn Reson Med*, 30(4):507–511, 1993.
- H. Lombaert, L. Grady, J.R. Polimeni, and F. Cheriet. Fast brain matching with spectral correspondence. In *Proc. IPMI*, pages 660–673. Springer, 2011.

- H. Lombaert, J. Sporring, and K. Siddiqi. Diffeomorphic spectral matching of cortical surfaces. In *Proc. IPMI*, pages 376–389. Springer, 2013.
- D.A. Low, M. Nystrom, E. Kalinin, P. Parikh, J.F. Dempsey, J.D. Bradley, S. Mutic, S.H. Wahab, T. Islam, G. Christensen, et al. A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing. *Med Phys.*, 30:1254–1263, 2003.
- D.A. Low, P.J. Parikh, W. Lu, J.F. Dempsey, S.H. Wahab, J.P. Hubenschmidt, M.M. Nystrom, M. Handoko, and J.D. Bradley. Novel breathing motion model for radiotherapy. *Int J Radiat Oncol.*, 63(3):921–929, 2005.
- L.J.P. van der Maaten, E.O. Postma, and H.J. Van Den Herik. Dimensionality reduction: A comparative review. *J Mach Learn Res.*, 10:1–41, 2009.
- D. Mah, J. Hanley, K.E. Rosenzweig, E. Yorke, L. Braban, C.C. Ling, S.A. Leibel, and G. Mageras. Technical aspects of the deep inspiration breath-hold technique in the treatment of thoracic cancer. *Int J Radiat Oncol*, 48(4):1175–1185, 2000.
- D. Manke, K. Nehrke, P. Börnert, P. Rösch, and O. Dössel. Respiratory motion in coronary magnetic resonance angiography: a comparison of different motion models. *J Magn Reson Imaging.*, 15(6):661–671, 2002.
- D. Manke, K. Nehrke, and P. Börnert. Novel prospective respiratory motion correction approach for free-breathing coronary MR angiography using a patient-adapted affine motion model. *Magn Reson Med.*, 50(1):122–131, 2003.
- D. Mateus, F. Cuzzolin, R. Horaud, and E. Boyer. Articulated shape matching using locally linear embedding and orthogonal alignment. In *Proc. ICCV*, pages 1–8. IEEE, 2007.
- D. Mateus, R. Horaud, D. Knossow, F. Cuzzolin, and E. Boyer. Articulated shape matching using Laplacian eigenfunctions and unsupervised point registration. In *IEEE Proc. CVPR*, pages 1–8. IEEE, 2008.
- J.R. McClelland, J.M. Blackall, S. Tarte, A.C. Chandler, S. Hughes, S. Ahmad, D.B. Landau, and D.J. Hawkes. A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy. *Med Phys.*, 33:3348–3358, 2006.
- J.R. McClelland, D.J. Hawkes, T. Schaeffter, and A.P. King. Respiratory motion models: A review. *Med Image Anal.*, 17(1):19 – 42, 2013.
- D.W. McRobbie, E.A. Moore, M.J. Graves, and M.R. Prince. *MRI from Picture to Proton*. Cambridge university press, 2006.

- M. Modat, G.R. Ridgway, Z.A. Taylor, M. Lehmann, J. Barnes, D.J. Hawkes, N.C. Fox, and S. Ourselin. Fast free-form deformation using graphics processing units. *Comput Meth Prog Bio.*, 98(3):278–284, 2010.
- W.W. Moses. Fundamental limits of spatial resolution in PET. *Nucl Instrum Meth A*, 648: S236–S240, 2011.
- K.S. Nayak, J.M. Pauly, D.G. Nishimura, and B.S. Hu. Rapid ventricular assessment using real-time interactive multislice MRI. *Magn Reson Med.*, 45(3):371–375, 2001.
- S.A. Nehmeh, Y.E. Erdi, C.C. Ling, K.E. Rosenzweig, O.D. Squire, L.E. Braban, E. Ford, K. Sidhu, G.S. Mageras, S.M. Larson, et al. Effect of respiratory gating on reducing lung motion artifacts in PET imaging of lung cancer. *Med Phys.*, 43(7):876–881, 2002.
- S.A. Nehmeh, Y.E. Erdi, G.S.P. Meirelles, O. Squire, S.M. Larson, J.L. Humm, and H. Schöder. Deep-inspiration breath-hold PET/CT of the thorax. *J Nucl Med*, 48(1): 22–26, 2007.
- S.A. Nehmeh, Y.E. Erdi, et al. Respiratory motion in positron emission tomography/computed tomography: a review. In *Semin Nucl Med.*, volume 38, pages 167–176, 2008.
- K. Nehrke, P. Börnert, D. Manke, and J. C. Böck. Free-breathing cardiac MR imaging: Study of implications of respiratory motion - initial results. *Radiology*, 220(3):810–815, 2001.
- K-H. Nennung, K. Kollndorfer, V. Schöpf, D. Prayer, and G. Langs. Multi-subject manifold alignment of functional network structures via joint diagonalization. In *Proc. IPMI*, pages 462–473. Springer, 2015.
- K. Ohara, T. Okumura, M. Akisada, T. Inada, T. Mori, H. Yokota, and M.J.B. Calaguas. Irradiation synchronized with respiration gate. *Int J Radiat Oncol*, 17(4):853–857, 1989.
- A. Okada, T. Murakami, K. Mikami, H. Onishi, N. Tanigawa, T. Marukawa, and H. Nakamura. A case of hepatocellular carcinoma treated by mr-guided focused ultrasound ablation with respiratory gating. *Magn Reson Med Sci*, 5(3):167–171, 2006.
- M. Panayiotou, A.P. King, R.J. Housden, Y. Ma, M. Cooklin, M. O’Neill, J. Gill, C.A. Rinaldi, and K.S. Rhode. A statistical method for retrospective cardiac and respiratory motion gating of interventional cardiac x-ray images. *Med Phys*, 41(7):071901, 2014.
- J.C. Park, S.H. Park, J.H. Kim, S.M. Yoon, S.Y. Song, Z. Liu, B. Song, K. Kauwelo, M.J. Webster, A. Sandhu, et al. Liver motion during cone beam computed tomography guided stereotactic body radiation therapy. *Med. Phys*, 39(10):6431–6442, 2012.

- Y. Pei, F. Huang, F. Shi, and H. Zha. Unsupervised image matching based on manifold alignment. *IEEE Pattern Anal*, 34(8):1658–1664, 2012.
- D. Peressutti. *Motion Modelling for Respiratory Motion Estimation in Minimally Invasive Cardiac Interventions Using INtraprocedure Ultrasound Data*. PhD thesis, School of Medicine, King’s College London, United Kingdom, 2014.
- D. Peressutti, E.J. Rijkhorst, D.C. Barratt, G.P. Penney, and A.P. King. Estimating and resolving uncertainty in cardiac respiratory motion modelling. In *Proc. ISBI.*, pages 262–265. IEEE, 2012.
- D. Peressutti, G.P. Penney, R.J. Housden, C. Kolbitsch, A. Gomez, E-J. Rijkhorst, D.C. Barratt, K.S. Rhode, and A.P. King. A novel bayesian respiratory motion model to estimate and resolve uncertainty in image-guided cardiac interventions. *Med Image Anal*, 17(4):488–502, 2013.
- J.G. Pipe et al. Motion correction with PROPELLER MRI: application to head motion and free-breathing cardiac imaging. *Magn Reson Med*, 42(5):963–969, 1999.
- C. Plathow, S. Ley, C. Fink, M. Puderbach, W. Hosch, A. Schmähel, J. Debus, and H-U. Kauczor. Analysis of intrathoracic tumor mobility during whole breathing cycle by dynamic MRI. *Int J Radiat Oncol*, 59(4):952–959, 2004.
- I. Polycarpou, C. Tsoumpas, and P.K. Marsden. Statistical evaluation of PET motion correction methods using MR derived motion fields. In *IEEE Nucl Sci Conf R.*, pages 3579–3585. IEEE, 2011.
- C. Prieto, S. Uribe, R. Razavi, D. Atkinson, and T. Schaeffter. 3D undersampled golden-radial phase encoding for DCE-MRA using inherently regularized iterative SENSE. *Magn Res Med*, 64(2):514–526, 2010.
- F. Qiao, T. Pan, John W. Clark Jr., and O.R. Mawlawi. A motion-incorporated reconstruction method for gated PET studies. *Phys Med Biol.*, 51(15):3769, 2006.
- B.W. Raaymakers, J.J.W. Lagendijk, J. Overweg, J.G.M. Kok, A.J.E. Raaijmakers, E.M. Kerckhof, R.W. van der Put, I. Meijsing, S.P.M. Crijns, F. Benedosso, et al. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept. *Phys Med Biol*, 54(12): N229, 2009.
- A. Rao, G.I. Sanchez-Ortiz, R. Chandrashekhara, M. Lorenzo-Valdés, R. Mohiaddin, and D. Rueckert. Comparison of cardiac motion across subjects using non-rigid registration. In *Proc. MICCAI.*, pages 722–729. Springer, 2002.

- M. Ries, B.D. De Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magn Reson Med*, 64(6):1704–1712, 2010.
- E.-J. Rijkhorst, I. Rivens, G. Haar, D. Hawkes, and D. Barratt. Effects of respiratory liver motion on heating for gated and model-based motion-compensated high-intensity focused ultrasound ablation. In *Proc. MICCAI.*, pages 605–612. Springer-Verlag, 2011.
- M.D. Robson, P.D. Gatehouse, M. Bydder, and G.M. Bydder. Magnetic resonance: an introduction to ultrashort TE (UTE) imaging. *J Comput Assist Tomo.*, 27(6):825–846, 2003.
- M. Rosenblatt et al. Remarks on some nonparametric estimates of a density function. *Ann math stat.*, 27(3):832–837, 1956.
- S.T. Roweis and L.K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.
- D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE T Med Imaging.*, 18(8):712–721, 1999.
- C. Santelli, R. Nezafat, B. Goddu, W.J. Manning, J. Smink, S. Kozerke, and D.C. Peters. Respiratory bellows revisited for motion compensation: preliminary experience for cardiovascular MR. *Magn Reson Med*, 65(4):1097–1102, 2011.
- L.K. Saul and S.T. Roweis. Think globally, fit locally: unsupervised learning of low dimensional manifolds. *J Mach Learn Res*, 4:119–155, 2003.
- F. Savill, T. Schaeffter, and A.P. King. Assessment of input signal positioning for cardiac respiratory motion models during different breathing patterns. In *Proc. ISBI.*, pages 1698–1701. IEEE, 2011.
- A. Schweikard, G. Glosser, M. Bodduluri, M.J. Murphy, and J.R. Adler. Robotic motion compensation for respiratory movement during radiosurgery. *Comput Aided Surg.*, 5(4):263–277, 2000.
- A. Schweikard, H. Shiomi, and J. Adler. Respiration tracking in radiosurgery without fiducials. *Int. J Med Rob Comp.*, 1(2):19–27, 2005.
- A.D. Scott, J. Keegan, and D.N. Firmin. Motion in cardiovascular MR imaging. *Radiology*, 250(2):331–351, 2009.

- Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. van Herk, J.V. Lebesque, and K. Miyasaka. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int J Radiat Oncol.*, 53(4): 822–834, 2002.
- Y. Seppenwoolde, R.I. Berbeco, S. Nishioka, H. Shirato, and B. Heijmen. Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study. *Med Phys*, 34(7):2774–2784, 2007.
- Gregory C Sharp, Steve B Jiang, Shinichi Shimizu, and Hiroki Shirato. Prediction of respiratory tumour motion for real-time image-guided radiotherapy. *Physics in medicine and biology*, 49(3):425, 2004.
- J.T. Sharp, N.B. Goldberg, W.S. Druz, and J. Danon. Relative contributions of rib cage and abdomen to breathing in normal subjects. *Journal of applied physiology*, 39(4):608–618, 1975.
- M. von Siebenthal, G. Székely, U. Gamper, P. Boesiger, A. Lomax, and P. Cattin. 4D MR imaging of respiratory organ motion and its variability. *Phys Med Biol.*, 52(6):1547–1564, 2007a.
- M. von Siebenthal, G. Székely, A. Lomax, and P.C. Cattin. Inter-subject modelling of liver deformation during radiation therapy. In *Proc. MICCAI*, pages 659–666. Springer, 2007b.
- M. von Siebenthal, G. Székely, A.J. Lomax, and P.C. Cattin. Systematic errors in respiratory gating due to intrafraction deformations of the liver. *Med Phys.*, 34:3620–3629, 2007c.
- R. Souvenir, Q. Zhang, and R. Pless. Image manifold interpolation using free-form deformations. In *Proc. IEEE Image Proc.*, pages 1437–1440. IEEE, 2006.
- M.K. Stam, S.P.M. Crijs, B.A. Zonnenberg, M.M. Barendrecht, M. van Vulpen, J.J.W. Lagendijk, and B.W. Raaymakers. Navigators for motion detection during real-time MRI-guided radiotherapy. *Phys Med Biol.*, 57(21):6797–6805, 2012.
- M.K. Stehling, R. Turner, P. Mansfield, et al. Echo-planar imaging: magnetic resonance imaging in a fraction of a second. *Science*, 254(5028):43–50, 1991.
- Y. Suh, S. Dieterich, B. Cho, and P.J. Keall. An analysis of thoracic and abdominal tumour motion for stereotactic body radiotherapy patients. *Phys Med Biol*, 53:3623–3640, 2008.
- H. Sundar, A. Khamene, L. Yatziv, and C. Xu. Automatic image-based cardiac and respiratory cycle synchronization and gating of image sequences. In *Proc. MICCAI*, pages 381–388. Springer, 2009.

- M. Tanter, M. Pernot, J-F. Aubry, G. Montaldo, F. Marquet, and M. Fink. Compensating for bone interfaces and respiratory motion in high-intensity focused ultrasound. *Int J Hyperther.*, 23(2):141–151, 2007.
- J.B. Tenenbaum, V. De Silva, and J.C. Langford. A global geometric framework for non-linear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.
- K. Thielemans, S. Mustafovic, and C. Tsoumpas. STIR: software for tomographic image reconstruction release 2. In *IEEE Nucl Sci Conf R.*, volume 4, pages 2174–2176. IEEE, 2006.
- K. Thielemans, S. Rathore, F. Engbrant, and P. Razifar. Device-less gating for PET/CT using PCA. In *Proc. IEEE NSS/MIC.*, pages 3904–3910. IEEE, 2011.
- M. Toriki, A. Elgammal, and C.S. Lee. Learning a joint manifold representation from multiple data sets. In *Proc. IEEE ICPR.*, pages 1068–1071, 2010.
- C. Tsoumpas, J.E. Mackewn, P. Halsted, A.P. King, C. Buerger, J.J. Totman, T. Schaeffter, and P.K. Marsden. Simultaneous PET–MR acquisition and MR-derived motion fields for correction of non-rigid motion in PET. *Ann Nucl Med.*, 24(10):745–750, 2010.
- A. Tzifa, G.A. Krombach, N. Krämer, S. Krüger, A. Schütte, M. von Walter, T. Schaeffter, S. Qureshi, T. Krasemann, E. Rosenthal, et al. Magnetic resonance–guided cardiac interventions using magnetic resonance–compatible devices. *Circ Cardiovasc Interv.*, 3(6):585–592, 2010.
- M. Uecker, S. Zhang, D. Voit, A. Karaus, K-D. Merboldt, and J. Frahm. Real-time MRI at a resolution of 20 ms. *NMR Biomed*, 23(8):986–994, 2010.
- J. Vandemeulebroucke, S. Rit, J. Kybic, P. Clarysse, and D. Sarrut. Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. *Med Phys.*, 38(1):166–181, 2011.
- S.S. Vedam, P.J. Keall, V.R. Kini, H. Mostafavi, H.P. Shukla, and R. Mohan. Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Phys Med Biol.*, 48(1):45–62, 2002.
- D. Visvikis, F. Lamare, P. Bruyant, N. Bousson, and C. Cheze Le Rest. Respiratory motion in positron emission tomography for oncology applications: Problems and solutions. *Nucl Instrum Meth A*, 569(2):453–457, 2006.
- C. Wachinger and N. Navab. Manifold learning for multi-modal image registration. In *Proc BMVC*, pages 1–12. Citeseer, 2010.

- C. Wachinger and N. Navab. Entropy and Laplacian images: Structural representations for multi-modal registration. *Med Image Anal.*, 16(1):1–17, 2012.
- C. Wachinger, M. Yigitsoy, E.J. Rijkhorst, and N. Navab. Manifold learning for image-based breathing gating in ultrasound and MRI. *Med Image Anal.*, 16(4):806–818, 2011.
- G. Wagenknecht, H-J. Kaiser, F.M. Mottaghy, and H. Herzog. MRI for attenuation correction in PET: methods and challenges. *Magn Reson Mater Phy*, 26(1):99–113, 2013.
- C. Wang and S. Mahadevan. Manifold alignment using Procrustes analysis. In *Proc. ICML.*, 2008.
- C. Wang and S. Mahadevan. Manifold alignment without correspondence. In *Proc. IJCAI*, pages 1273–78, 2009.
- Y. Wang, S.J. Riederer, and R.L. Ehman. Respiratory motion of the heart: kinematics and the implications for the spatial resolution in coronary imaging. *Magn Reson Med.*, 33(5):713–719, 1995.
- M.J. White, D.J. Hawkes, A. Melbourne, D.J. Collins, C. Coolens, M. Hawkins, M.O. Leach, and D. Atkinson. Motion artifact correction in free-breathing abdominal MRI using overlapping partial samples to recover image deformations. *Magn Reson Med.*, 62(2):440–449, 2009.
- S. Wold, K. Esbensen, and P. Geladi. Principal component analysis. *Chemometr Intell Lab*, 2(1):37–52, 1987.
- C. Würslin, H. Schmidt, P. Martirosian, C. Brendle, A. Boss, N.F. Schwenzer, and L. Stegger. Respiratory motion correction in oncologic PET using T1-weighted MR imaging on a simultaneous whole-body PET/MR system. *J Nucl Med.*, 54(3):464–471, 2013.
- T. Yamamoto, U. Langner, B.W. Loo, J. Shen, and P.J. Keall. Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients. *Int J Radiat Oncol*, 72(4):1250–1258, 2008.
- C. Yao, J. Simpson, T. Schaeffter, and G. Penney. Multi-view 3D echocardiography compounding based on feature consistency. *Phys Med Biol.*, 56(18):6109, 2011.
- H. Zaidi and A. Del Guerra. An outlook on future design of hybrid PET/MRI systems. *Med Phys.*, 38(10):5667–5689, 2011.

- C.J. Zech, K.A. Herrmann, A. Huber, O. Dietrich, A. Stemmer, P. Herzog, M.F. Reiser, and S.O. Schoenberg. High-resolution MR-imaging of the liver with T2-weighted sequences using integrated parallel imaging: Comparison of prospective motion correction and respiratory triggering. *J Magn Reson Imaging*, 20(3):443–450, 2004.
- D. Zhai, B. Li, H. Chang, S. Shan, X. Chen, and W. Gao. Manifold alignment via corresponding projections. In *Proc. BMVC.*, pages 3–11, 2010.
- Q. Zhang, R. Souvenir, and R. Pless. On manifold structure of cardiac MRI data: Application to segmentation. In *IEEE Proc. CVPR.*, volume 1, pages 1092–1098. IEEE, 2006.
- Q. Zhang, Y.C. Hu, F. Liu, K. Goodman, K.E. Rosenzweig, and G.S. Mageras. Correction of motion artifacts in cone-beam CT using a patient-specific respiratory motion model. *Med Phys.*, 37(6):2901–2909, 2010.