

Modelling and Reconstructing the Respiratory Motion of the Liver

Inauguraldissertation

Zur Erlangung der Würde eines

Dr. sc. med.

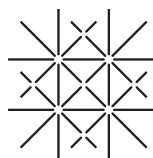
vorgelegt der Medizinischen Fakultät der Universität Basel

von

Frank Preiswerk

aus Basel

Basel, 2013



UNI
BASEL

Genehmigt von der Medizinischen Fakultät auf Antrag von

Prof. Dr. Philippe C. Cattin, Basel
Dissertationsbetreuer

Prof. Dr. Antony Lomax, Villigen
Koreferent

Prof. Dr. med. Daniel Oertli, Basel
Externer Gutachter

Basel, den

Prof. Dr. med. Christoph Beglinger
Dekan

In memory of

Emmanuel Esser
and
Max Kammermann

who introduced me to Engineering and Computer Science.

Acknowledgements

I would like to acknowledge my supervisor Prof. Dr. Philippe C. Cattin for giving me the opportunity to pursue a PhD in his group. It is inspiring and fun to work with somebody so creative, focused, efficient and appreciative. His support was remarkably hands-on regarding the complete spectrum of my work. I would also like to acknowledge my fellow PhD student and close collaborator Dr. Patrik Arnold, whose sharp analytical mind was invaluable for this work. Dr. Beat Fasel deserves acknowledgement for keeping our infrastructure up and running, including our high performance computation cluster in spite of his own research projects. Furthermore, I would like to acknowledge my collaborators in our SNF project. These are (in alphabetical order): Prof. Dr. Oliver Bieri, Zarko Celicanin, Prof. Dr. Antony Lomax, Valeria De Luca, Lorena Petrusca, Dr. Rares Salomir, Dr. Francesco Santini, Prof. Dr. Klaus Scheffler and Dr. Christine Tanner. In addition I also owe thanks to Dr. Reinhard Knothe for his advice. A big thank you goes to all members of our lab for the many insightful discussions about research and other things during the past months and years.

Finally, I want to thank my parents for their generous support during my studies.

Summary

Respiratory organ motion is a complicating factor in tumour treatment. For dose delivery, the aim is to obtain a possibly static target in the beam's eye view whenever the beam is on. Such a reduction of motion allows for reducing the safety margins or for delivering a higher dose in a shorter time. This thesis proposes a statistical, population-based model that covers all modes of deformation such as the perpetual breathing motion and organ drift. It furthermore provides a mathematical tool to estimate the current organ position based on sparse and low-dimensional measurements, with the goal to use ultrasound for obtaining the respiratory signal directly from the liver. The model can describe complex variations of the liver in shape and position without knowing the underlying physical mechanisms. To achieve this, 4D-MRI is acquired for a number of subjects. From these images, deformation fields are extracted, inter-subject correspondence is established and the model is learned from the data. The prediction accuracy is evaluated in various simulations where partial information of the organ in 3d, 2d or only 1d is known. Furthermore, an experiment is described where simultaneous 4D-MRI and ultrasound is acquired for six subjects in order to evaluate the approach in a clinically relevant scenario. The prediction is driven by tracked points in the ultrasound images and then compared to the ground-truth obtained from 4D-MRI. The results show that a statistical motion model can significantly reduce the uncertainty with respect to organ position during respiration.

Zusammenfassung

Die Atembewegung von Organen ist ein erschwerender Faktor in der Tumorbehandlung. In der Dosisplanung ist es wünschenswert, die zu bestrahlende Region als möglichst statisches Ziel im Fokus zu halten. Eine Reduktion der Bewegung aus der Sicht des Bestrahlungsgerätes erlaubt es, Sicherheitsmargen zu verkleinern sowie höhere Dosen in kürzerer Zeit zu verabreichen. In dieser Arbeit wird ein statistisches, populations-basiertes Modell vorgeschlagen, welches alle für die Strahlentherapie relevanten Arten von Organbewegung wie die zyklische Atembewegung sowie Organdrift abdeckt. Ausserdem wird ein mathematisches Verfahren vorgestellt, um die momentane Organposition von lückenhaften und niedrig-dimensionalen Messwerten zu rekonstruieren. Das Modell ist in der Lage, die komplexe Veränderung der Leber in Form und Position zu beschreiben, ohne Kenntnis der zugrundeliegenden physikalischen Mechanismen zu haben. Um dies zu erreichen, werden 4D-MRI Sequenzen der Leber von zahlreichen Personen aufgenommen. Von diesen Bildern werden anschliessend die Bewegungsfelder extrahiert sowie Korrespondenz zwischen den verschiedenen Lebern etabliert, um das Modell zu lernen. Die Genauigkeit der Vorhersage wird in verschiedenen Simulationen evaluiert. Im Weiteren wird ein Experiment beschrieben, in dem von sechs Personen 4D-MRI und Ultraschall gleichzeitig aufgezeichnet wurde, um den Ansatz in einem klinisch relevanten Szenario zu testen. Dabei wurde die Vorhersage mittels verfolgten Punkten im Ultraschall berechnet und anschliessend mit der tatsächlichen Bewegung in 4D-MRI verglichen. Die Resultate zeigen, dass ein statistisches Bewegungsmodell die Ungenauigkeit in Bezug auf die Organposition während der Atmung massgebend reduzieren kann.

Contents

Contents	v
Nomenclature and Abbreviations	viii
1 Introduction	1
1.1 Physiology of Respiratory Motion	2
1.2 Types of Cancer Treatments	5
1.2.1 Invasive Methods	5
1.2.2 Minimally-Invasive Methods	6
1.3 Managing Respiratory Motion in Radiation Therapy	8
1.4 Contribution of this Thesis	11
2 Data Acquisition	13
2.1 4D-MRI	13
2.2 Image Registration	15
3 Statistical Motion Modelling	17
3.1 PCA Model	18
3.1.1 Properties of the PCA	20
3.1.2 Establishment of Correspondence	21
3.1.3 Model Overview	25
3.2 Reconstruction from Sparse Data	28
3.2.1 Bayes' Theory	28
3.2.2 Bayesian Approach to Reconstruction	29
3.3 Publications	32

CONTENTS

4	A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements	33
5	Robust Tumour Tracking From 2D Imaging Using a Population-Based Statistical Motion Model	43
6	Towards More Precise, Minimally-Invasive Tumour Treatment Under Free Breathing	51
7	Model-Guided Respiratory Organ Motion Prediction of the Liver from 2d Ultrasound	57
8	Conclusions	71
	References	73

Nomenclature

Mathematical Notation

s	subject
n_s	total number of subjects
n_s^t	total number of time steps of s
$n = \sum_s n_s^t$	total number of training samples
m	number of 3d model points
t	time
$\mathbf{v} \in \mathbb{R}^{3m}$	data sample (surface and interior)
$\check{\mathbf{v}} \in \mathbb{R}^{3m}$	data sample at exhalation
$\check{\mathbf{v}}^\circ \in \mathbb{R}^{3m'}$	data sample surface at exhalation
$\mathbf{x} = \mathbf{v} - \check{\mathbf{v}} \in \mathbb{R}^{3m}$	shape-free data sample (motion vector)
$\boldsymbol{\mu} \in \mathbb{R}^{3m}$	mean vector of all shape-free training samples
$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{3m \times n}$	matrix of training samples
$\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_{n-1}] \in \mathbb{R}^{3m \times n-1}$	orthonormal matrix of $n - 1$ eigenvectors
$\mathbf{c} = (c_1, \dots, c_{n-1})^T$	vector of $n - 1$ model coefficients
T	translation matrix
R	rotation matrix

Nomenclature

Medical Imaging Modalities

MRI	magnetic resonance imaging
4D-MRI	4-dimensional MRI (3d + time)
CT	computed tomography
US	ultrasound

Anatomical Directional Terms

S-I	superior-inferior
A-P	anterior-posterior
L-R	left-right
supine	body lying down with the face up

Cancer Treatment Terms

IMRT	intensity-modulated radiotherapy
IMPT	intensity-modulated proton therapy
HIFU	high-intensity focused ultrasound
IGRT	image-guided radiotherapy
RFA	radio-frequency ablation

Chapter 1

Introduction

Cancer is a leading cause of death worldwide. According to the most recent cancer report of the World Health Organization (WHO) [1], it is the second leading cause of death in developed countries and among the three leading causes of death for adults in developing countries. In 2008, there were an estimated 12.3 million incident cases of cancer worldwide and 7.6 million deaths of cancer (13% of all deaths). Knowledge about the prevention and treatment of cancer is increasing, yet the number of cases grows every year. The dominant types of cancer are lung, stomach, liver, colon and breast tumours. These sites are subject to respiratory motion, which can be a complicating factor for treatment. Radiotherapy is a well known and established method for the treatment of non-resectable tumours and over 50% of all cancer patients will require radiotherapy at some time during their illness [2]. The challenge in radiotherapy is to deliver a lethal dose to all cancerous tissue without destroying healthy cells. There are a number of different factors contributing to potential inaccuracies in irradiation. Of these, respiratory motion is one of the major problems in the thorax and in the abdomen and has been shown to have a large dosimetric impact on conventional radiotherapy [3, 4, 5]. Besides quasi-periodic respiratory motion, which includes variation in breathing depth and speed, the organs undergo other modes of deformation, called secondary modes [6]. The secondary modes are caused, for example, by the cardiac cycle motion, digestive activity, gravity, muscle relaxation, or filling of the bladder. These two components constitute the total organ motion seen during treatment. Thanks to the advances in image-guided

radiotherapy (IGRT) in recent years that paved the way for technologies like intensity-modulated radiotherapy (IMRT), intensity-modulated proton therapy (IMPT) [7] and high-frequency focused ultrasound (HIFU) [8, 9, 10], it is possible to deliver highly precise dose distributions even to complex tumour shapes near vitally important structures.

To take advantage of the higher precision of these novel tumour treatment modalities, methods to cope with uncertainties due to respiratory motion are desired. In this thesis, techniques are developed for real-time tumour location prediction of the liver, based on the position of multiple tracked landmarks, *e.g.* the diaphragm or vessel branches, in ultrasound images. 4D-MRI images [11] are acquired that allow to measure the 3d organ motion over longer periods of time. This 4d motion data is used to build and evaluate mathematical models for abdominal organs, in our case the liver, and also serves as ground-truth to evaluate the accuracy in various simulations. To evaluate the approach in a clinical setup, simultaneous 4D-MRI/ultrasound organ motion data is acquired using a custom, MR-compatible ultrasound transducer. The fiducials are extracted from the ultrasound images and the prediction is compared to the ground-truth obtained from 4D-MRI through non-rigid registration.

A central concern in this work is to avoid invasive technologies as much as possible. With the exception of a pre-operative 3d CT scan of the subject, which is currently standard procedure in clinical practice and can be replaced by an MRI scan, all methods proposed in this thesis are completely non-invasive.

1.1 Physiology of Respiratory Motion

Respiration is a quasi-periodic, bio-mechanical process that consists of an inhalation phase followed by an exhalation phase. The most dominant muscle involved in this process is the diaphragm, to which the inferior end of the lung is attached. During inhalation, the diaphragm contracts in order to expand the lung volume by pulling it in inferior direction. The intercostal muscles connect adjacent ribs and also participate in inhalation. They contract during inhalation, pulling the ribs superiorly and anteriorly, thereby increasing both the lateral and anterior-posterior diameters of the thorax. This allows air to enter the lung, where gas

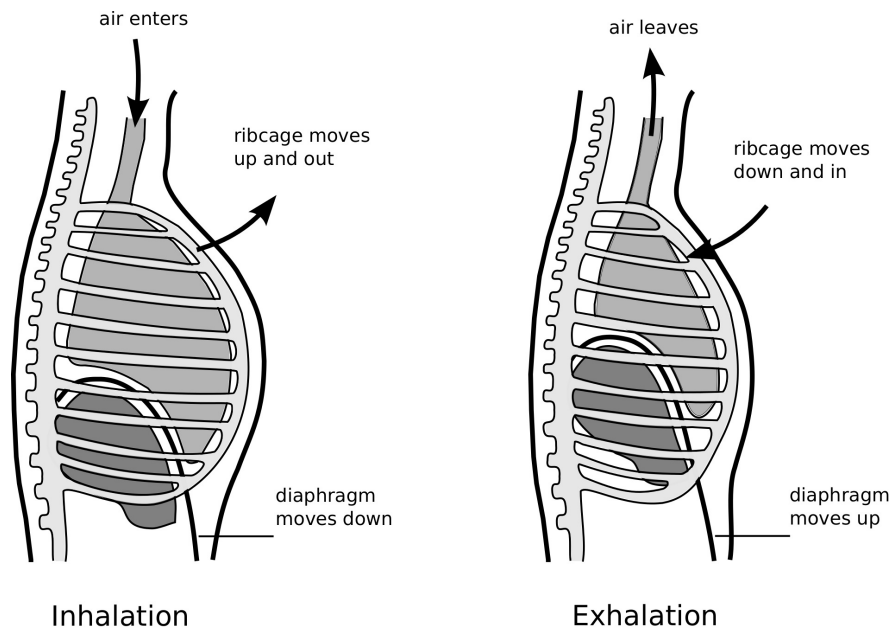


Figure 1.1: Physiology of respiratory motion within the thorax.

exchange takes place between blood and air within the capillaries. Exhalation is a passive process, where the diaphragm relaxes and both lung and rib cage, due to their elasticity, move back in superior and posterior inferior direction, respectively. Figure 1.1 depicts the respiratory motion of the thorax for inhalation and exhalation. This form of breathing, where respiration is dominantly caused by contraction of the diaphragm, is called *abdominal* breathing. On the other hand, *costal* breathing is a rare type of breathing that occurs when negative intrapulmonary pressure is primarily achieved by contraction of the external intercostal muscle. A comprehensive overview on the physiology of respiratory motion can be found in [12].

The respiratory process also affects organs not directly involved in respiration. Many other organs are known to move with respiration, as summarised in an overview study by the American Association of Physicists in Medicine (AAPM) [13], see Tab. 1.1. The diaphragm pushes the abdominal organs, *e.g.* the liver, in inferior direction. During exhalation, they return to approximately the same position as at the beginning of the cycle. While the average *superior-inferior* (S-I) motion for abdominal organs is anywhere from 10 mm to 25 mm, the average

Site	Observer	Breathing mode	
		Shallow [mm]	Deep [mm]
Pancreas	Suramo <i>et al.</i> [17]	20 (10-30)	43 (20-80)
	Bryan <i>et al.</i> [18]	20 (0-35)	-
Liver	Weiss <i>et al.</i> [19]	13 +/-5	-
	Harauz <i>et al.</i> [20]	14	-
	Suramo <i>et al.</i> [17]	25 (10-40)	55 - (30-80)
	Davies <i>et al.</i> [14]	20 (10-30)	
Kidney	Suramo [17] <i>et al.</i>		
	Davies <i>et al.</i> [14]		
Diaphragm	Wade [21]	17	101
	Korin <i>et al.</i> [15]	13	39
	Davies <i>et al.</i> [14]	12 (7-28)	43 (25-57)
	Weiss <i>et al.</i> [19]	13 +/- 5	-
	Giraud <i>et al.</i> [22]	-	35 (3-95)
	Ford <i>et al.</i> [23]	50 (13-31)	-

Table 1.1: Extent of abdominal organ motion in superior-inferior (S-I) direction. The mean range of motion and the (minimum–maximum) or its std. dev (+/-) for each site and each cohort of subjects (between 10 and 30). Overview of different studies from [13].

anterior-posterior (A-P) and *left-right* (L-R) motion is no more than 2 mm [14, 15], although for the liver, A-P motion has been found to be larger than that in another study [16].

Variation of Respiratory Motion Although breathing motion often follows a repetitive pattern, it is generally irregular. The breathing motion for each cycle is caused by contributions from the diaphragm and the intercostal muscles that may vary over time. Naturally, the position of a person, *e.g.* whether someone is positioned supine or upright, also affects breathing.

In the literature, two main types of breathing variation are distinguished. *Intra-cycle variation* describes the variation within one breathing cycle, not the same for inhalation as for exhalation and *inter-cycle variation* is the variation between one breathing cycle to another cycle. One particular type of intra-cycle variation is *hysteresis*, describing a separation between the trajectories during

inhalation and exhalation of up to 5 mm [24]. Additionally, in the field of radiotherapy and medical imaging, the variation between different treatment or imaging sessions (*fractions*) is relevant: *Intra-fraction variation* is the variation of motion within one fraction, while *inter-fraction variation* is the variation of motion among different fractions, typically over days or weeks.

1.2 Types of Cancer Treatments

This chapter gives a brief overview on the different types of cancer treatment used today. On a high level, the approaches can be divided into invasive and minimally-invasive methods, both of which have their advantages and drawbacks. This chapter describes the relevant techniques for both cases..

1.2.1 Invasive Methods

Surgically removing a tumour is the preferred form of treatment if position, shape and size of the tumour allow it [25]. Surgery is very efficient if the entire malignancy can be removed, as it causes comparatively little side effects besides the ordeal of the surgery itself. However, sometimes surgery is not possible, *e.g.* if the tumour is located near important vessels or other vital structures. In such cases, only parts of the tumour, if at all, can be surgically removed and the remaining cancerous tissue must be treated by other means.

Chemotherapy is a form of drug therapy and the first choice in a number of malignancies [26]. It is also often performed in addition to surgery in cases where a certain risk exists that a tumour might start to grow again. However, there is a series of side-effects that render chemotherapy highly inconvenient for any patient. They include depression of the immune system, nausea, vomiting and loss of hair. For smaller tumours, another alternative to surgery is radio-frequency ablation (RFA) [27, 28]. It destroys the malignant tissue by heating it up using high-frequency alternative current. Therefore, a needle that contains multiple curved, retractable electrodes is inserted into the tumour by the oncologist. Image guidance is necessary to correctly place the needle inside the tumour. RFA is used for both small primary tumours as well as for metastases. A problem is the risk

of mechanical disruption and spreading of tumor cells into normal tissues due to the radiofrequency needle [29].

Despite the high efficiency of surgery and the aforementioned treatment methods in cases where they fit in well, it is desirable to treat tumours non-invasively, *i.e.* using only external means. In some cases it is even the only option. For the remainder of this chapter, we will therefore focus on external therapy methods, which are relevant in this context.

1.2.2 Minimally-Invasive Methods

There are number of treatment methods if surgery is not an option or to complement both surgery and and chemotherapy. The goal in radiotherapy is to kill malignant cells using ionising radiation, which destroys DNA of the target cells leading to cellular death using either photons (gamma-rays or x-rays), charged particles (electrons, protons or heavy ions) or uncharged particles (neutrons). External beam radiotherapy (EBRT) encompasses all techniques where the radiation is delivered to the tumour from outside of the body. A problem is that photons and particles interact with all tissue, not just the tumour, depositing varying levels of energy as they pass through the body. While the maximum dosage can normally be delivered to the target location (although this depends on the depth), there is a significant amount of dosage delivered to the surrounding tissue using conventional radiotherapy. To circumvent this limitation, intensity-modulated radiotherapy (IMRT) [30] irradiates the tumour from various angles using photons, allowing to reduce the dosage delivered to the healthy tissue between radiation source and tumour as well as beyond. It also allows to deliver dosage to more complex target volume shapes. However, the radiation still damages healthy cells outside the tumour and the resulting damage increases the probability of secondary cancer [31, 32].

In particle therapy (PT), protons, neutrons or positive ions are used for radiation, whereas proton therapy is the most common form. Protons and ions have the advantage that energy can be deposited much more sharply at specific depths within the body compared to conventional radiotherapy. Like in radiotherapy, the DNA is damaged which leads to cell death. Figure 1.2 compares the typical

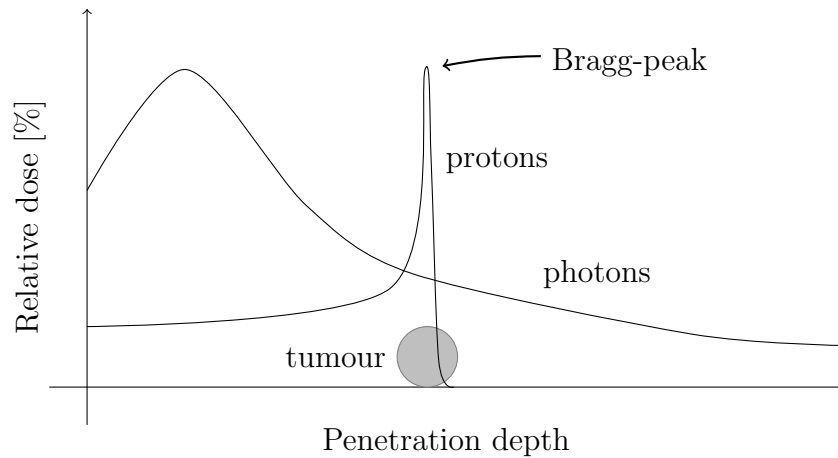


Figure 1.2: Energy-deposition curves for protons and photons.

absorption curves of photons and protons. The sharp peak at the target depth for protons is called *Bragg-peak*. Intensity-modulated proton therapy (IMPT) adapts to more complex shapes and additionally minimises dosage to healthy tissue similar to IMRT [7]. Whilst very promising results have been observed with proton therapy, one problem is the size and cost of the necessary synchrotron for particle acceleration.

In neutron therapy [33], a delivery agent containing a non-radioactive isotope (typically boron-10) is injected into the patient. As neutrons pass through the body, they interact with the agent, resulting in a biologically destructive reaction. Clinical interest in neutron therapy is mainly on the treatment of high grade gliomas, recurrent cancers of the head and neck region and either primary or metastatic melanoma. The main problem that remains today is how to design a delivery agent that is more tumour-selective .

In internal radiotherapy, or brachytherapy [34], radiation is administered from within the tissue. To treat prostate cancer, for example, radioactive seeds (containing iodine-125, palladium-103 or iridium-194) are inserted into the prostate gland. They are either permanently left inside (permanent low dose radiation, LDR) or removed shortly after (high dose radiation, HDR).

Another promising treatment technique that is currently being researched is high-intensity focused ultrasound (HIFU) [8, 9, 10]. Using a phased array of ultrasound transducers, a focal spot can be created using constructive interference

that can be steered in 3d. The sonication with powerful sound waves results in an increase in temperature to a certain point where denaturation of the abnormal tissue occurs. There are certain advantages compared to radiation- and proton therapy. First, the method is non-ionising. Furthermore, such a setup is much simpler, cheaper and more space saving than for example a proton therapy system. Another advantage is the ability to control the focal spot electronically without mechanical parts involved, which can also be space-saving. Finally, such a device can be made MRI compatible, thus temperature mapping can be used during treatment to measure and control tissue heating and therefore safety margins. Disadvantages include longer treatment times, limited depth of penetration and the absorption of sound waves in cases of overlying ribs or lung as well as a lack of long-time studies.

A central issue that is common to all of the above approaches (with the exception of neutron therapy) is tumour motion. In contrast to *e.g.* the brain, abdominal organs cannot be fixated during treatment. The advantage of highly localised treatment methods makes these techniques sensitive to respiratory organ motion and drift, which represents a limiting factor for exploiting their full potential. The possibility of depositing sharply localised energy within the body necessitates the use of more precise motion compensation systems for moving organs, for no or bad motion compensation would only damage healthy tissue with "higher precision" while delivering no dosage to the tumour itself in the worst case.

1.3 Managing Respiratory Motion in Radiation Therapy

A lot of research has been done on handling organ motion during the past ten years. Comprehensive overviews on the research field can be found in [35] and [36] and practical guidelines have been published in a recent report of the American Association of Physicists in Medicine [13].

The goal of respiratory motion management is to keep the target and the treatment beam aligned throughout the entire breathing cycle. More precise

knowledge about the position of the target is necessary to make use of the improved precision in dose delivery with IMPT or HIFU as well as to improve conventional radiotherapy.

A simple method to avoid respiratory motion is to completely interrupt breathing while the therapy beam is on [37, 38, 39, 40, 41]. This method, called *breath-hold*, has the obvious disadvantage that the treatment time is prolonged and some individuals, particularly lung patients, might not be able to hold their breath at all. Another straight-forward approach is to reduce the extent of breathing by asking the patient to breath shallowly [42, 43, 44, 45], which is often not sufficient on its own and therefore combined with other methods. In *gated* treatment, the beam is only turned on during a certain fraction of the breathing phase, typically near exhalation [4, 46, 47, 48, 49]. Although the aforementioned approaches compensate breathing motion to some extent, they require reproducibility of the organ position for the selected breathing phase [6] and prolong the treatment time. More importantly, they only compensate for the perpetual breathing motion and as such are completely oblivious to all other modes of organ motion. Thus they are only accurate in a short window of a couple of minutes after patient set-up. Hence, organ motion during radiotherapy continues to be a problem and much research effort is being put into understanding and addressing this issue.

It would be desirable to keep the target and the treatment beam aligned throughout the entire breathing cycle. This technique is commonly referred to as *tracking*. Tracking is very demanding as it requires some kind of knowledge about the target motion [50, 51, 52]. Although tracking is in principle designed to follow any target motion, it profits from a possibly regular breathing pattern because this simplifies the required short-time prediction of the motion trajectory. Any tracking method must be fed with some kind of surrogate signal upon which an estimate on the position of the organ can be made. Such surrogate signals can be external or internal. A spirometer that measures the air flow using a mouth piece provides a simple external surrogate signal [53]. Measuring the displacement of the abdominal skin using a respiration belt or by optical tracking are other options [47]. Although these methods provide a temporally high-resolved respiratory signal, it has been shown that the motion of organs does not necessarily correlate well with such external signals [54, 55, 56], either due to organ drift or because

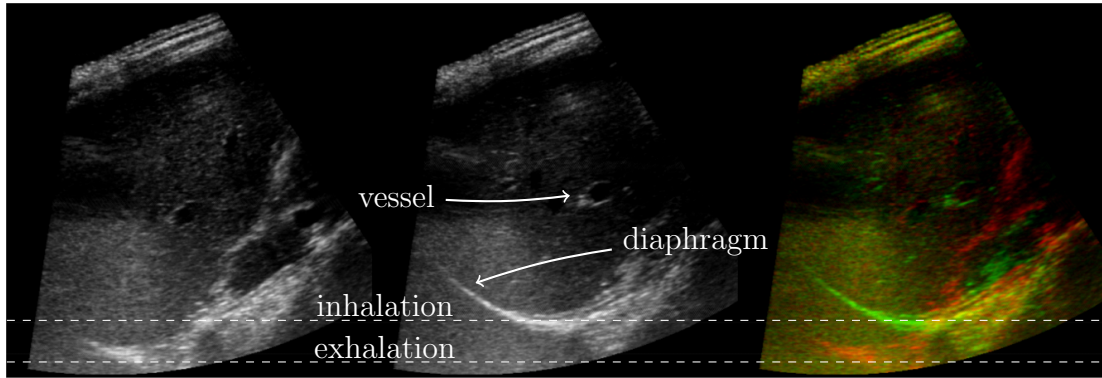


Figure 1.3: Ultrasound as a modality for image-guided therapy. The above images depict a liver at exhalation (left) and at inhalation (middle). The diaphragm is clearly visible in both images and is therefore suitable for tracking the respiratory state. Additionally, several liver vessels are visible and can be used for tracking as well. The right image shows an overlay of both images.

of varying motion patterns at different positions of the organ. It has also been observed that organs may drift significantly over time due to effects like intestinal activity, gravity or relaxation [6], thereby invalidating the established relation between the exhalation position and the treatment device.

In image-guided radiotherapy (IGRT), imaging techniques are used to increase the accuracy of target localisation. Fluoroscopy has been used in clinical interventions for many years, despite continuous exposure of the subject to radiation. Portal imaging is a technique where the treatment device is at the same time used as an imaging device [57], but the contrast of such MV-images is very low. Moreover, radiation-based imaging techniques stand for even higher exposure of the patient. A very attractive non-ionising imaging modality for IGRT is ultrasound (US). Its high sampling rate and low lag make it suitable to use for tracking. Figure 1.3 depicts a liver in ultrasound at both inhalation and exhalation. The diaphragm is clearly visible in both phases. However, only few tumours are visible under US and it completely fails for anything hidden behind absorbing structures like the ribs or inside the lung. Thus, conventional US alone is rather limited in its capability as a tumour tracking modality, as the tumour can be invisible on the ultrasound and only surrogate information can be observed. On the other hand, the theoretical investigation on the liver, reported in [11], showed that the knowledge about the position of one or more surrogate 3d points (not necessarily the tumour directly) and a mathematical model of the organ of interest allows to predict the position of the entire organ with high spatial precision.

More specifically, tracking can serve two distinct purposes. With known surrogate positions at time t , we distinguish between:

Spatial prediction: Reconstruction of the complete organ position based on partial information available from the surrogate at time t .

Temporal prediction: Temporal prediction of the surrogate position for $t + \Delta t$ from the known location at time t .

Spatial prediction is used to compute an estimate of a specific target position when only a surrogate of the actual target is available. Temporal prediction, on the other hand, is about the time lag occurring because of the necessary processing of the tracking signal and adjusting the beam in a real-world situation. A precise treatment system must therefore predict a certain amount of time into the future. By combining the two, *spatio-temporal* prediction of the complete organ for $t + \Delta t$ can be achieved, so in practise they are often intertwined. The focus of this thesis is on spatial prediction, however, in Chapter 7 we also present approaches for temporal prediction that can be easily combined with our reconstruction methods, thereby providing a system for the complete motion compensation problem as a whole.

1.4 Contribution of this Thesis

In this thesis, techniques are presented for robust real-time organ motion prediction of the liver based on sparse information of one or more, possibly low-dimensional fiducials. A Bayesian formulation is employed to reconstruct the complete shape of the liver at any time based on the surrogate data. The approach is based on a population-based mathematical motion model built from 4D-MRI data of multiple subjects. The work of von Siebenthal *et al.* [16] is closely related to the developments in this thesis. They built three types of motion models for the liver:

- Type I Subject-specific model of drift and respiratory motion
- Type II Population-based model of anatomical shape variation (at exhalation)

- Type III Population-based model of inter-fraction drift (at exhalation)

In that work, it is mentioned that subject-specific motion data might not be possible to acquire for a subject in clinical practice, given the long acquisition times necessary. Another prohibitive factor not mentioned is the cost of MRI acquisitions. Therefore, the author focused on population-based models (types II and III) and showed that drift can be predicted for an unseen subject based on a PCA model of previously recorded data.

The methods developed in this thesis are a combination and extension of the above approaches. We propose a population-based statistical motion model where no subject-specific motion data is necessary for prediction. Our goal is to investigate on how accurate the complete motion of an unknown liver can be reconstructed from partial information and a previously obtained population-based respiratory motion model.

The model is built from 4D-MRI images of 20 subjects. This data also serves as the ground-truth to evaluate the accuracy in various leave-one-out experiments. We describe application scenarios where information is given in 3d (*e.g.* implanted electromagnetic beacons, implanted gold markers or stereo kV-imaging [57, 58, 59]), 2d (*e.g.* tracked points in ultrasound or portal images [60]) or 1d (*e.g.* internal or external respiratory signal [53]). The ultimate goal is to use the model in conjunction with structures tracked in 2d US. This is done for six subjects by acquiring 4D-MRI and US images simultaneously using MR-compatible equipment specifically designed by our project collaborators. From this data, the accuracy can be validated by comparing the reconstructions based on the US images to the true motion observed in 4D-MRI.

For establishing correspondence between a new subject and our model, a single MRI scan is sufficient. Consequently, all methods proposed in this thesis are potentially non-invasive. However, acquiring a pre-operative 3d CT scan of a patient is standard procedure for dose calculation in radiotherapy. Such a scan can also be used for fitting the model.

Chapter 2

Data Acquisition

The motion models described in this thesis are obtained from 4D-MRI sequences of up to 20 subjects in supine position. An excellent overview on 4D-MRI can be found in [16]. This chapter briefly summarises the key concepts.

2.1 4D-MRI

Magnetic resonance imaging (MRI) is a non-ionising image modality that provides good contrast between soft tissues, which makes it useful for imaging abdominal organs and cancerous tissue. It is based on a powerful magnetic field to align the magnetisation of hydrogen atoms in the body. This magnetisation is then systematically altered through radio-frequency pulses. As the protons return to their original state (aligned to the magnetic field), a radio-frequency signal is emitted, which can be measured by a receiver coil. Spatial information can be extracted by applying a magnetic gradient, causing nuclei at different locations to precess at different speeds. Like in computed tomography (CT), only an image of a single slice can be acquired at a time (however, in arbitrary orientation and thickness). Volumetric images are obtained by stacking multiple slices together. For non-moving structures, this can be done with high spatial resolution at the cost of a long image acquisition. However, in the case of moving organs, the acquisition time for each slice must be as short as possible to avoid motion artefacts. Furthermore, in the case of respiratory motion, images acquired at different

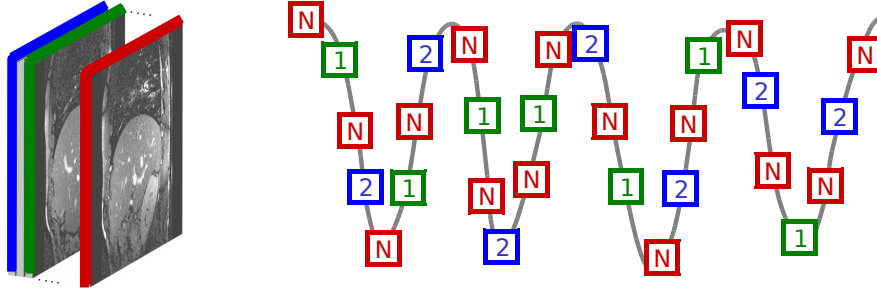


Figure 2.1: The 4D-MRI acquisition scheme. Each slice position $(1, 2, \dots)$ is temporally encapsulated by two navigators (N). The navigator slice is always acquired at the same spatial position and defines a phase for each encapsulated image. This allows to find matching slices of the same breathing phase for each position and thus gives a 3d stack for every point in time.

times generally correspond to different phases of the motion cycle. This renders the reconstruction of correct volumetric data non-trivial. 4D-MRI is based on the assumption that anatomically correct 3d images can be obtained by collecting 2d images of the same respiratory state from different respiratory cycles. To achieve this, data slices of the organ at equidistant spatial positions that cover the complete organ are repeatedly acquired, typically in sagittal direction. One complete sweep of the organ is called a *dynamic*. Because the organ is subject to respiratory motion throughout the acquisition of a dynamic, each of its data slices corresponds to a different respiratory phase. By acquiring a large number of dynamics, the probability increases that for every position a data slice is obtained at every phase of a (typical) complete respiratory cycle. Additionally, *navigator* images at constant position are acquired in an interleaved manner after each data slice. From the navigator pair that embraces each data slice, a respiratory phase can be assigned to it. When enough data has been acquired, a 3d image can be retrospectively reconstructed for each encapsulating navigator image pair, thereby creating a 4d image sequence. Figure 2.1 depicts the 4D-MRI acquisition scheme.

2.2 Image Registration

The motion data used to learn our motion model is obtained from the 4D-MRI data through non-rigid registration, which gives 4D deformation fields of each subject (see Fig. 2.2). There is a trade-off between the degrees of freedom of a registration algorithm (*i.e.* the irregularity of deformation it can describe) and the smoothness of the resulting deformation field. We focused our evaluation of registration methods on the Demons algorithm [61] and the B-Spline method [62]. There were two main criteria for our application: The transformation model of the registration method should be anatomically accurate and second it should not be distracted by discontinuities occurring at the border of the organ. Experiments on our data revealed that while the Demons algorithm is in principle able to describe complex deformations, it falls significantly short in terms of the anatomical correctness of the resulting field. The algorithm turns out to be rather susceptible to overfitting due to its large number of degrees of freedom, particularly with the kind of images dealt with in this context. The B-Spline algorithm on the other hand computes a deformation of only a small number of regular grid points and its inherent regularisation property leads to smoother, anatomically reasonable deformation fields. The negative side is that it approximates discontinuous motion with smooth motion fields at the organ border. Therefore we opt for the approach described in [16] where the B-Spline method is used on a masked moving image that is registered to unmasked images. The approach follows a multi-resolution scheme and uses normalised mutual information (NMI) as a similarity measure [63, 64]. The control point resolution is 15 mm isotropically.

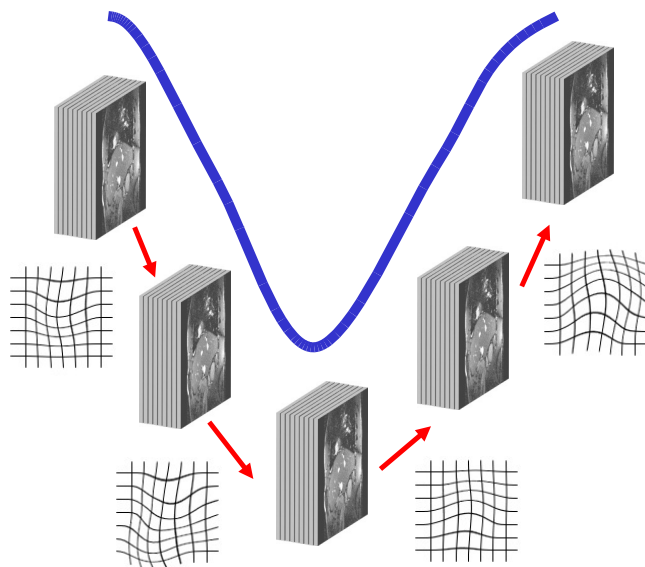


Figure 2.2: Non-rigid registration of the 4D-MRI stacks results in dense spatio-temporal motionfields.

Chapter 3

Statistical Motion Modelling

Statistical models have been heavily used for shape modelling in the past 20 years. Cootes *et al.* captured the invariant properties of a class of 2d shapes using a Point Distribution Model (PDM) [65, 66]. Every object is represented using a set of labelled points. The mean position of the points represents an average shape and a number of modes of variation are determined describing the main modes in which the example shapes tend to deform from the average. Principal component analysis (PCA) is the mathematical tool used to compute the model. This idea was later extended to 3d objects, most notably in the context of face reconstruction [67]. In [16], a statistical model is proposed to model subject-specific respiratory motion and population-based drift of the liver. In this chapter we explain how motion is modelled using PCA and unfold the methods used for spatial prediction from sparse measurements.

In a general sense, a model is a description of a system using mathematical concepts. Models lie at the heart of many algorithms in machine learning and biomedical image analysis, as they are a means to compactly represent (anatomical) knowledge about a specific domain. In learning theory, for example, a typical problem is to predict the class label $y \in \{-1, +1\}$ from an observation x , whose distribution is generally unknown (but often assumed to be normal). A *generative* model is able to describe instances of the modelled class by learning a model of joint probability $p(x, y)$ between samples x and class labels y , *i.e.* by modelling $p(y = +1|x)$ and $p(y = -1|x)$. In contrast, a discriminant model approach only models the posterior $p(y|x)$ directly. Thus, discriminative models, as the name

suggests, focus on the classification boundaries between the positives and negatives, whereas generative models emphasise the data generation process in each individual class [68]. Taking the concept of generative models to the domain of unsupervised learning, a model that allows to generate new samples of the underlying distribution by learning the distribution $p(x)$ of a set of hidden variables x is a generative model. Since a model instance is defined by the parameter x , it is also a *parametric* model.

Following the terminology in [69], we define

Definition A *linear generative statistical model* is a model

$$\mathcal{M} : (\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathbf{U}) \quad (3.1)$$

where $\boldsymbol{\mu} \in \mathbb{R}^{3m}$ is the mean, $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_n]$ is a matrix with n components and σ is a weighting of the components, with the following properties:

- Given a coefficient vector \mathbf{c} , a model instance can be generated by the mean plus a linear combination of modes

$$\mathbf{x} = \boldsymbol{\mu} + \sum_i c_i \sigma_i \mathbf{u}_i. \quad (3.2)$$

- The coefficients

$$c_i \sim \mathcal{N}(0, 1) \quad (3.3)$$

are independent and distributed according to a normal distribution.

3.1 PCA Model

Principal component analysis (PCA) is a method that transforms a set of input data into a representation where each dimension is independent and distributed according to a normal distribution. The PCA model

$$\mathcal{M}_{PCA} : (\boldsymbol{\mu}_{PCA}, \boldsymbol{\sigma}_{PCA}, \mathbf{U}_{PCA}) \quad (3.4)$$

is one particular type of model after to the above definition. We will drop the subscript indices for simplicity from now on as we work exclusively with PCA models. Let $\mathbf{x}_i = (x_{i,1}, y_{i,1}, z_{i,1}, \dots, x_{i,m}, y_{i,m}, z_{i,m})^T$, $i \in 1, \dots, n$ represent vectors of 3d data points arranged in a single column. The arithmetic mean is estimated using

$$\boldsymbol{\mu} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i. \quad (3.5)$$

The matrix $\mathbf{X} = [\mathbf{x}_1 - \boldsymbol{\mu}, \dots, \mathbf{x}_n - \boldsymbol{\mu}] \in \mathbb{R}^{3m \times n}$ is formed by arranging all mean-free data vectors. Let further $\Sigma \in \mathbb{R}^{3m \times 3m}$ be the covariance matrix of \mathbf{X} defined by

$$\Sigma = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T. \quad (3.6)$$

PCA diagonalises the covariance matrix using singular value decomposition (SVD),

$$\Sigma = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T. \quad (3.7)$$

such that $\mathbf{U} \in \mathbb{R}^{3m \times n-1}$ is an orthonormal basis of principal components of \mathbf{X} and the diagonal matrix $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_{n-1})$ contains the variance along the dimensions in PCA space. Computing this SVD of Σ becomes increasingly expensive with high-dimensional data. Fortunately, there is a simple relation between Eq. 3.7 and the SVD of the data matrix \mathbf{X} ,

$$\Sigma = \frac{1}{n} \mathbf{X}\mathbf{X}^T \quad (3.8)$$

$$= \frac{1}{n} \mathbf{U}\mathbf{W}\mathbf{V}^T(\mathbf{U}\mathbf{W}\mathbf{V})^T \quad (3.9)$$

$$= \frac{1}{n} \mathbf{U}\mathbf{W}\mathbf{V}^T\mathbf{V}^T\mathbf{W}\mathbf{U}^T \quad (3.10)$$

$$= \frac{1}{n} \mathbf{U}\mathbf{W}^2\mathbf{U}^T \quad (3.11)$$

$$= \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T, \quad (3.12)$$

where $\boldsymbol{\Lambda} = \frac{1}{n} \mathbf{W}^2 \in \mathbb{R}^{n-1 \times n-1}$. This equality allows to obtain the SVD of Σ directly from $\mathbf{X} = \mathbf{U}\mathbf{W}\mathbf{V}^T$. From the diagonal entries $\mathbf{W} = \text{diag}(w_1, \dots, w_n)$ we

can further obtain the *standard deviation* of the transformed data,

$$\sigma_i = \sqrt{\lambda_i} = \frac{w_i}{\sqrt{n}}. \quad (3.13)$$

It is common to sort the columns of \mathbf{U} in descending order according to the magnitude of the corresponding eigenvalues λ_i . Furthermore, the PCA projection can be scaled with the inverse of the standard deviation, resulting in the following transformation of a data vector into PCA space,

$$\mathbf{c} = \text{diag}(\sigma_1^{-1}, \dots, \sigma_{n-1}^{-1}) \mathbf{U}^T (\mathbf{x} - \boldsymbol{\mu}). \quad (3.14)$$

According to Eq. 3.2, the back-projection is then defined by

$$\mathbf{x} = \boldsymbol{\mu} + \sum_i c_i \sigma_i \mathbf{u}_i = \boldsymbol{\mu} + \mathbf{U} \mathbf{c}. \quad (3.15)$$

3.1.1 Properties of the PCA

The PCA domain has a number of useful properties with respect to the distribution of the data. The rescaling according to Eq. 3.14 yields coefficients $c_i \sim \mathcal{N}(0, 1)$ that are not only *independent* but also *identically* distributed. This property will be important later for reconstruction from sparse data.

PCA models the data based on a multivariate normal distribution

$$p(\mathbf{x}) = \nu_c \prod \exp\left(-\frac{1}{2\sigma_i^2} \langle (\mathbf{x} - \boldsymbol{\mu}), \mathbf{u}_i \rangle^2\right) \quad (3.16)$$

$$= \nu_c \cdot \exp\left(-\frac{1}{2} \sum \frac{1}{\sigma_i^2} \langle (\mathbf{x} - \boldsymbol{\mu}), \mathbf{u}_i \rangle^2\right) \quad (3.17)$$

$$= \nu_c \cdot \exp\left(-\frac{1}{2} \mathbf{c}^T \mathbf{c}\right) = \nu_c \cdot \exp\left(-\frac{1}{2} \|\mathbf{c}\|^2\right) \quad (3.18)$$

with the normalisation constant $\nu_c = ((2\pi)^{k/2} \cdot |\text{diag}(\sigma)|^{1/2})^{-1}$, but when the coefficients are rescaled to unit variance, it reduces to $\nu_c = (2\pi)^{k/2}$ and the probability density for \mathbf{c} can be expressed as simply

$$p(\mathbf{c}) = \nu_c \cdot \exp\left(-\frac{1}{2} \|\mathbf{c}\|^2\right). \quad (3.19)$$

Another property is the inherent compression ability of PCA. The dimensionality of the data can be reduced by discarding the k smallest eigenvalues λ_i , which leads to coefficients

$$\tilde{\mathbf{c}} = \text{diag}(\sigma_i)_k^{-1} \mathbf{U}_k^T (\mathbf{x} - \boldsymbol{\mu}), \quad (3.20)$$

with $\mathbf{U}_k = [\mathbf{u}_1, \dots, \mathbf{u}_{n-1-k}]$ and $\text{diag}(\sigma_i)_k^{-1} = \text{diag}(\sigma_1^{-1}, \dots, \sigma_{n-1-k}^{-1})$. Let $\tilde{\mathbf{C}} = [\tilde{\mathbf{c}}_1, \dots, \tilde{\mathbf{c}}_{n-1-k}]$ be the matrix of compressed coefficients of the data matrix \mathbf{X} . The reconstructed data matrix

$$\tilde{\mathbf{X}} = \text{diag}(\sigma_i)_k \mathbf{U}_k \tilde{\mathbf{C}} + \mathbf{M}, \quad \mathbf{M} = [\boldsymbol{\mu}, \dots, \boldsymbol{\mu}] \in \mathbb{R}^{3m \times n} \quad (3.21)$$

is optimal with respect to the squared distance $\mathbf{e}_i = \|\mathbf{x}_i - \tilde{\mathbf{x}}_i\|^2$ between each original data point and its approximation, averaged over the entire data set,

$$E = \frac{1}{n} \sum_{i=1}^n \mathbf{e}_i = \frac{1}{n} \sum_{i=1}^n \|\mathbf{x}_i - \tilde{\mathbf{x}}_i\|^2. \quad (3.22)$$

The sum of the eigenvalues corresponding to the retained eigenvectors is a measure for the total variance contained in the compressed data set. Selecting between 95% and 99% variance is common to remove dimensions that are redundant or only encode noise.

3.1.2 Establishment of Correspondence

In order to learn the statistical motion model described in Section 3.1, correspondence must be established among all data samples. In other words, a common topology must be defined for the data to be modelled. The reason for this is that we want the motion pattern of a specific position in the liver to influence only the statistics of exactly this position. Let the vector $\mathbf{v}_s(t)$, $s \in 1, \dots, n_s^t$ describe the liver of subject s at time t (with points located on the surface and within the liver). The dependency on s and t hints that we must distinguish between two types of correspondence: *intra-subject* and *inter-subject* correspondence. Intra-subject correspondence is associated with all time steps t for a single subject, while inter-subject correspondence concerns the establishment of a common data format among all subjects.

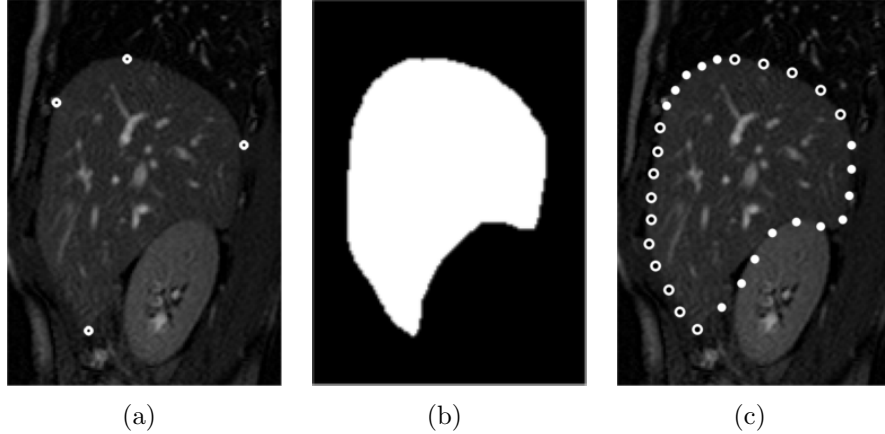


Figure 3.1: (a) Manually selected landmarks for correspondence on a selected sagittal slice. (b) Manually defined mask. (c) Automatically generated resampling of the liver contour based on mask and landmarks. The alternating markers highlight the four individually controlled segments.

Inter-Subject Correspondence We want to establish mechanical correspondence [16], which means that the correspondence criterion should be primarily based on how similar points move during respiration. As elaborated in Sec. 1.1, this motion is driven by the diaphragm and guided by the abdominal wall as a “tube” defining the direction where the organ can move. We therefore select four landmark points on each slice of a master exhalation 4D-MRI image for each subject: the anterior and posterior points of the liver where the diaphragm touches the abdominal wall, the most superior point adjacent to the diaphragm and the most inferior point. Figure 3.1(a) shows one slice with these four landmarks.

Additionally, the liver is segmented manually on every slice of this master exhalation image (Fig. 3.1(b)). Although it would be tempting to use one of many available methods that could in principle automate this task, due to the high speed of image acquisition in 4D-MRI the image quality is not sufficient to reliably automate liver segmentation in our case. Particularly, it is not our goal to trade-off precision for processing speed because we want to introduce as little error as possible in our model. After the liver has been segmented, the previously selected landmarks are projected onto the segmented liver contour. The four resulting contour segments are subsampled equidistantly to give the

final set of 30 points per segment. This process is repeated for all slices. Figure 3.1(c) shows the result on our exemplar slice. For each of the segment points $i \in 1, \dots, 30$ we define a B-spline controlled by the corresponding point on all slices $(p_{i,1}, p_{i,2}, \dots, p_{i,k})$, that is, in out- of-plane (left- right) direction, where k is the number of slices. The spline is then resampled at 30 locations. This is done because image acquisitions of different subjects generally consist of different numbers of slices. After resampling the splines, the approach is independent of k . For each subject, we get a vector of corresponding surface points at exhalation, $\check{\mathbf{v}}_s^\circ$.

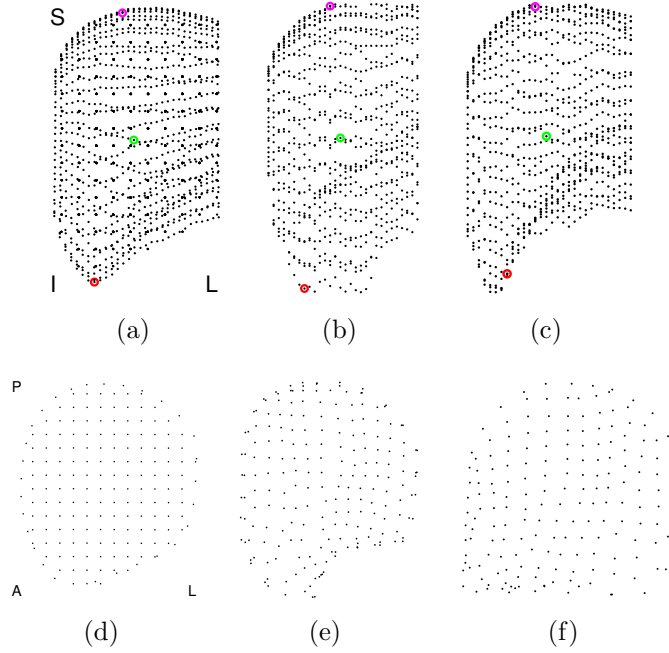


Figure 3.2: (a) Mean exhalation shape with isotropic grid and three points highlighted. (b),(c) Master exhalation shape of two selected subjects with the same corresponding points highlighted. (d)-(f) Slice through shape that shows internal grid points for the above examples.

Up to this point, the intra-subject correspondence of the *surface* is defined and the next step is to define the correspondence of the *interior*. Therefore we first align all master shapes $\check{\mathbf{v}}_s^\circ$ using partial (no scaling) generalised procrustes analysis (GPA) [70] to compute a rigid transformation T_s, R_s for each subject into

a common coordinate system. For each subject s , the method finds an optimal transform according to

$$\arg \min_{T_s, R_s} \sum_{s=1}^n \|T_s R_s \tilde{\mathbf{v}}_s^\circ - \boldsymbol{\mu}^\circ\|_2, \quad (3.23)$$

where T is a translation matrix, R_s an orthogonal rotation matrix and $\boldsymbol{\mu}^\circ = \frac{1}{n} \sum_s \mathbf{v}_s^\circ$ is the mean surface. The “goodness-of-fit” criterion is the l_2 -norm. Procrustes alignment is an iterative procedure that is usually repeated until convergence. The algorithm outline is the following:

1. Arbitrarily choose a reference shape among all shapes
2. Superimpose all shapes $\tilde{\mathbf{v}}_s^\circ$ to current reference shape
3. Compute the mean shape $\boldsymbol{\mu}^\circ$ of the current set of superimposed shapes
4. Compute the l_2 -norm between the reference and all shapes, set reference to mean shape $\boldsymbol{\mu}^\circ$ and continue to step 2

This is repeated a fixed number of times or until the change of the l_2 -norm between the old and the new reference in step 4 falls below a threshold. The algorithm’s output is a set of transformations (T_s, R_s) for each subject and the mean shape $\boldsymbol{\mu}^\circ$.

In order to establish *interior* correspondence, $\boldsymbol{\mu}^\circ$ is embedded within an isotropic grid of 10 mm spatial resolution. The grid points are warped to each subject $\tilde{\mathbf{v}}_s^\circ$ based on the position of their surface points and a tetrahedrisation of the grid points [71, 72]. This way, the mechanical correspondence is transferred from the surface to the inside of the liver and an exhalation master shape $\tilde{\mathbf{v}}_s$ with points on the surface as well as in the interior is obtained for each subject. Figure 3.2 depicts the final correspondence points on the mean shape as well as on two subjects.

Intra-Subject Correspondence Recall that correspondence is linked to both the subjects s and to time t . The latter means that *temporal* correspondence must be established along the motion sequence for each subject. However, given the inter-subject correspondence from the previous section, this is now as simple

as warping the correspondence points of the master shape $\check{\mathbf{v}}_s$ to all individual time steps using the deformation field obtained from the non-rigid registration. Using this approach, the location of every point in the master shape is known over time. In other words, dense motion information $\mathbf{v}_s(t)$ of the liver is obtained for each subject.

3.1.3 Model Overview

For the final computation of the PCA model, a number of cycles from the beginning, middle and end of each acquisition session, are selected, as the inclusion of all states would only increase the computational cost with minimal information gain. Since we are interested in modelling motion, not shape, only the offsets between each state and the exhalation master are included in the model,

$$\mathbf{x}_s(t) = \mathbf{v}_s(t) - \check{\mathbf{v}}_s, \quad (3.24)$$

and the data is processed according to Sec. 3.1. For all experiments in this thesis, motion models built from between 12 and 20 subjects are used. Figure 3.3 depicts coronal and sagittal views of the exhalation masters from six subjects. Figure 3.4 shows the vertex displacement of one respiratory cycle of a selected model vertex and the first two PCA coefficients of the associated complete liver. It can be seen that, since PCA is a linear model, the overall shape of the trajectory between the vertex displacement and the first two PCA coefficients is similar, although the direction of the trajectory is generally not the same in the two spaces. This is because there are two solutions in an eigenvalue decomposition that differ in the sign of the resulting eigenvalue (and its associated eigenvector), although the solutions are equivalent. Which eigenvalue is chosen depends on the SVD algorithm employed. Intra-cycle variation (hysteresis) is visible between inhalation and exhalation in both plots for most subjects. This shows that the registration algorithm used is susceptible enough to discover smaller variations in the data, despite the rather coarse image resolution of 4D-MRI. Furthermore, the cyclic pattern is extracted correctly by the algorithm, as cycle start and end are close both in the spatial domain as well as in PCA space.

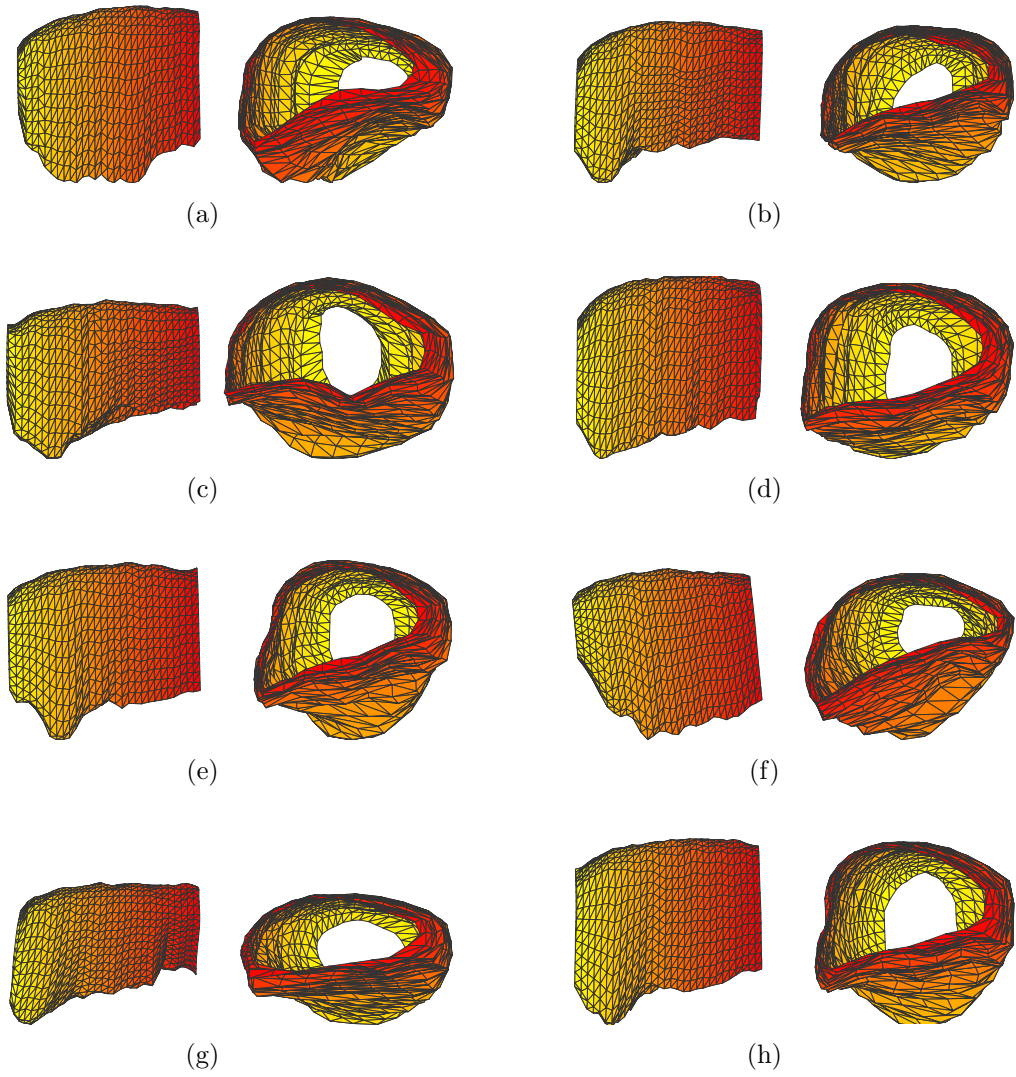


Figure 3.3: Coronal (left) and sagittal (right) view of the surface of six exhalation master shapes $\check{\mathbf{v}}_s$.

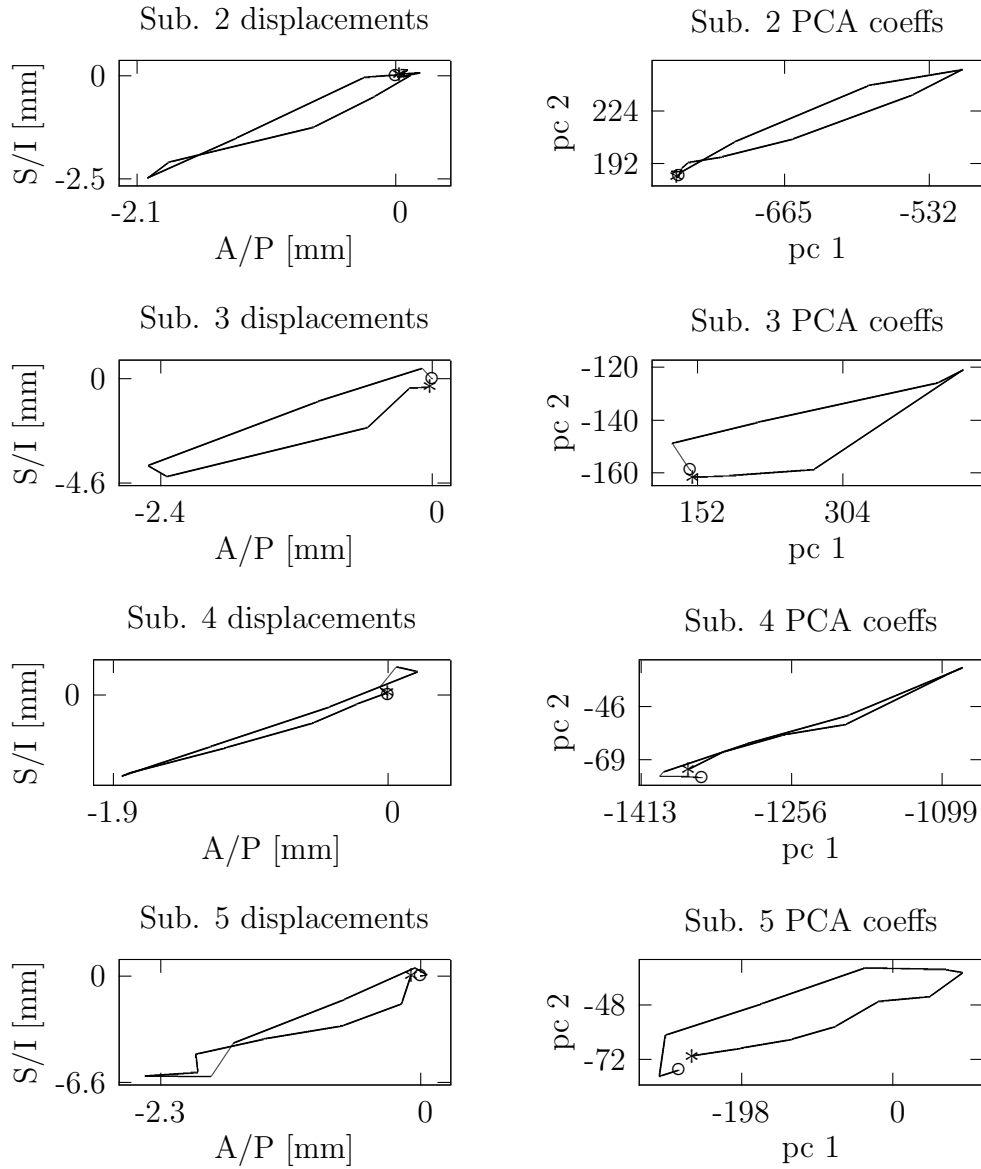


Figure 3.4: Vertex displacement (left) and the first and second PCA coefficients (right) of one respiratory cycle. The asterisk (*) highlights the start of the cycle and the circle (o) highlights its end.

3.2 Reconstruction from Sparse Data

The probabilistic framework of PCA provides an efficient tool to reconstruct missing information based on the statistics of the modelled data. In order to explain how this is done, we make use of Bayes' theory, which will be recapped shortly, before the reconstruction algorithm is described.

3.2.1 Bayes' Theory

Let X and Y be two random variables. The two fundamental rules of probability are the *sum rule* and the *product rule*:

$$\text{sum rule} \quad p(X) = \sum_Y p(X, Y) \quad (3.25)$$

$$\text{product rule} \quad p(X, Y) = p(Y|X)p(X), \quad (3.26)$$

where $p(X, Y)$ is the joint probability of X and Y and $p(Y|X)p(X)$ is the conditional probability of Y given X . Bayes' rule relates the two conditional probabilities $P(X|Y)$ and $P(Y|X)$ and is of fundamental importance in many machine learning applications. It is derived from the product rule and the symmetry property $p(X, Y) = p(Y, X)$:

$$p(X|Y) = \frac{p(Y|X)p(X)}{p(Y)}. \quad (3.27)$$

Using the sum rule, the denominator in Eq. 3.27 can be expressed with the quantities appearing in the numerator,

$$p(Y) = \sum_X p(Y|X)p(X). \quad (3.28)$$

In machine learning, X often takes the form of an observation vector \mathbf{x} and Y takes the form of a set of model parameters Θ . The conditional probability

$$p(\mathbf{x}|\Theta) \quad (3.29)$$

is called the *likelihood* of \mathbf{x} given a set of model parameters Θ . It describes how probable an observed data vector is for different settings of the parameter vector Θ . The value $p(\Theta)$ gives the probability of the model parameters *before* a sample has yet been drawn and is called *prior probability* of Θ . Accordingly, the quantity $p(\Theta|\mathbf{x})$ gives the probability of a set of model parameters *after* a sample \mathbf{x} has been drawn and is therefore called *posterior probability* of Θ .

3.2.2 Bayesian Approach to Reconstruction

The goal in model-based reconstruction is to obtain an estimation of the complete data vector \mathbf{x} from a vector of partially observed information $\mathbf{r} \in \mathbb{R}^l, l < 3m$. We want to make use of the statistical properties of the PCA to obtain a reasonable estimation based on the assumption of normal distribution and Bayes' theory. Based on the previously described set of rules, we can now understand the Bayesian approach in [73] that was used for reconstruction.

Let $\mathbf{L} : \mathbb{R}^{3m} \mapsto \mathbb{R}^l, l < 3m$ be a linear transformation matrix that governs the mapping of a complete data vector to its partial observation,

$$\mathbf{r} = \mathbf{L}\mathbf{x}. \quad (3.30)$$

In general, \mathbf{L} is not an injective mapping and thus the solution of Eq. 3.30 with respect to \mathbf{x} is not uniquely defined. Likewise, we can define the reduced version of the model matrix \mathbf{U} , scaled with the eigenvalues,

$$\mathbf{Q} = (\mathbf{q}_1, \dots, \mathbf{q}_n) = \text{diag}(\sigma_i) \cdot \mathbf{L}\mathbf{U} \in \mathbb{R}^{l \times (n-1-k)}. \quad (3.31)$$

For the projection of an element that is defined by the model coefficient vector \mathbf{c} , a noiseless measurement would be

$$\hat{\mathbf{r}} = \mathbf{Q}\mathbf{c}. \quad (3.32)$$

If we assume that the measurement is subject to uncorrelated Gaussian noise of

variance σ_N^2 , the likelihood of observing $\mathbf{r} \in \mathbb{R}^l$ is given by

$$p(\mathbf{r} | \dot{\mathbf{r}}) = \prod_{j=1}^l P(r_j | \dot{r}_j) \quad (3.33)$$

$$= \prod_{j=1}^l \nu \cdot \exp\left(-\frac{1}{2\sigma_N^2}(\dot{r}_j - r_j)^2\right) \quad (3.34)$$

$$= \nu_N^l \cdot \exp\left(-\frac{1}{2\sigma_N^2} \sum_j (\dot{r}_j - r_j^2)\right) \quad (3.35)$$

$$= \nu_N^l \cdot \exp\left(-\frac{1}{2\sigma_N^2} \|\dot{\mathbf{r}} - \mathbf{r}\|^2\right) \quad (3.36)$$

with a normalisation factor $\nu_N = ((2\pi)^{l/2} \cdot \sqrt{\sigma_N})^{-1}$. Equation 3.36 can be expressed in terms of the model coefficient,

$$p(\mathbf{r} | \mathbf{c}) = \nu_N^l \cdot \exp\left(-\frac{1}{2\sigma_N^2} (\|\mathbf{Q}\mathbf{c} - \mathbf{r}\|^2)\right). \quad (3.37)$$

At this point, we can make use of Bayes' rule (Eq. 3.27) to compute $P(\mathbf{c} | \mathbf{r})$, the posterior probability of a model parameter \mathbf{c} given a partial observation \mathbf{r} ,

$$p(\mathbf{c} | \mathbf{r}) = \frac{p(\mathbf{r} | \mathbf{c}) \cdot p(\mathbf{c})}{p(\mathbf{r})}. \quad (3.38)$$

Our ultimate goal is to optimise this probability. According to Eq. 3.28, the prior probability $p(\mathbf{r})$ in the denominator is another constant factor $\nu = (\sum_{\mathbf{c}'} P(\mathbf{r} | \mathbf{c}') p(\mathbf{c}'))^{-1}$. To find a solution, we substitute Eq. 3.19 and 3.37 into the posterior probability in Equation 3.38,

$$p(\mathbf{c} | \mathbf{r}) = \nu \cdot \nu_N^l \cdot \nu_c \cdot \exp\left(-\frac{1}{2\sigma_N^2} (\|\mathbf{Q}\mathbf{c} - \mathbf{r}\|^2)\right). \quad (3.39)$$

Taking the logarithm of 3.39 yields an energy function that has to be minimised,

$$E(\mathbf{c}) = -2 \cdot \log p(\mathbf{c} | \mathbf{r}) = \frac{1}{\sigma_N^2} (\|\mathbf{Q}\mathbf{c} - \mathbf{r}\|^2) + \|\mathbf{c}\|^2 + \text{const.} \quad (3.40)$$

We introduce a regularisation parameter $\eta = \sigma_N^2$ and minimise

$$E(\mathbf{c}) = \|\mathbf{Q}\mathbf{c} - \mathbf{r}\|^2 + \eta \cdot \|\mathbf{c}\|^2. \quad (3.41)$$

Expansion of this energy function gives

$$E(\mathbf{c}) = \langle \mathbf{Q}\mathbf{c}, \mathbf{Q}\mathbf{c} \rangle - 2\langle \mathbf{Q}\mathbf{c}, \mathbf{r} \rangle + \|\mathbf{r}\|^2 + \eta \cdot \|\mathbf{c}\|^2. \quad (3.42)$$

Since $\langle \mathbf{Q}\mathbf{c}, \mathbf{Q}\mathbf{c} \rangle = (\mathbf{Q}\mathbf{c})^T \cdot \mathbf{Q}\mathbf{c} = \mathbf{c}^T \cdot \mathbf{Q}^T \mathbf{Q}\mathbf{c} = \langle \mathbf{c}, \mathbf{Q}^T \mathbf{Q}\mathbf{c} \rangle$, and similar for $\langle \mathbf{Q}\mathbf{c}, \mathbf{r} \rangle$, the optimum is

$$0 = \nabla E(\mathbf{c}) = 2\mathbf{Q}^T \mathbf{Q}\mathbf{c} - 2\mathbf{Q}^T \mathbf{r} + 2\eta\mathbf{c} \quad (3.43)$$

which leads to the following equation to be solved for \mathbf{c} ,

$$\mathbf{Q}^T \mathbf{Q}\mathbf{c} + \eta\mathbf{c} = \mathbf{Q}^T \mathbf{r}. \quad (3.44)$$

After \mathbf{Q} is substituted with its singular value decomposition $\mathbf{Q} = \hat{\mathbf{U}}\mathbf{W}\hat{\mathbf{V}}^T$, we obtain

$$\hat{\mathbf{V}}\mathbf{W}^2\hat{\mathbf{V}}^T\mathbf{c} + \eta\mathbf{c} = \hat{\mathbf{V}}\mathbf{W}\hat{\mathbf{U}}^T\mathbf{r}. \quad (3.45)$$

The following series of transformations leads to the final solution,

$$\mathbf{W}^2\hat{\mathbf{V}}^T\mathbf{c} + \eta\hat{\mathbf{V}}^T\mathbf{c} = \mathbf{W}\hat{\mathbf{U}}^T\mathbf{r} \quad (3.46)$$

$$\text{diag}(w_i^2 + \eta) \cdot \hat{\mathbf{V}}^T\mathbf{c} = \mathbf{W}\hat{\mathbf{U}}^T\mathbf{r} \quad (3.47)$$

$$\hat{\mathbf{V}}^T\mathbf{c} = \text{diag}\left(\frac{w_i}{w_i^2 + \eta}\right)\hat{\mathbf{U}}^T\mathbf{r} \quad (3.48)$$

$$\mathbf{c} = \hat{\mathbf{V}}\text{diag}\left(\frac{w_i}{w_i^2 + \eta}\right)\hat{\mathbf{U}}^T\mathbf{r}. \quad (3.49)$$

The final estimation of the complete data vector is computed according to Eq. 3.15,

$$\mathbf{x} = \mathbf{U}\mathbf{c} + \mu. \quad (3.50)$$

Since only motion information is contained in our model, the exhalation master shape must be added to obtain the absolute position of the organ, $\mathbf{v} = \mathbf{x} + \check{\mathbf{v}}$.

The matrix \mathbf{Q} is of size $l \times \hat{m}$, with $\hat{m} \ll m$ the number of principal components used for reconstruction. Therefore, Eq. (3.49) can be solved in real-time.

Note that the above described reconstruction approach was introduced for reconstructing the 3d shape of a face from a set of landmarks, its texture and a population-based morphable model of the human face [73]. There, the available evaluation criterions are primarily qualitative, as a typical test could be how realistic a face looks. In our case, precision is of much higher importance. However, as we will later see, it turns out that the approach is in fact suitable for the reconstruction of organ motion under tight quantitative criterions.

3.3 Publications

The next sections present the peer-reviewed publications carried out in the context of this thesis. In all studies, the above described population-based statistical motion model and the Bayesian reconstruction algorithm are used. The chapters follow their chronological order of publication and start off with the introduction of the model and rather theoretical simulations where a number of assumptions are made. Then, these assumptions are gradually loosened over the course of two additional publications and finish with a comprehensive study carried out in a clinical setup, where ultrasound images are used for tracking the organ and simultaneously acquired 4D-MRI is used for validation. The latter is a joint development together with our project partners.

Chapter 4

A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements

This paper, presented at the *Workshop on Computational and Clinical Applications in Abdominal Imaging (ABD-MICCAI)*, September 2011, Toronto, Canada¹, introduces the basics of our population-based statistical motion model. It is assumed that accurate 3d-information of three markers within the liver is given. Based on these surrogates, the position of the complete organ is predicted using our Bayesian approach. Leave-one-out experiments are performed for a model built from 4D-MRI data of 12 subjects. The average error for each subject is between 0.8 mm and 2 mm and the average error over all experiments is 1.2 mm.

¹The article is available online at http://link.springer.com/content/pdf/10.1007/2F978-3-642-28557-8_26.

A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements

Frank Preiswerk, Patrik Arnold, Beat Fasel, and Philippe C. Cattin

Medical Image Analysis Center, University of Basel, Switzerland
{frank.preiswerk,philippe.cattin}@unibas.ch

Abstract. In this paper, we present an approach for modelling and predicting organ motion from partial information. We used 4D-MRI sequences of 12 subjects to build a statistical population model for respiratory motion of the liver. Using a Bayesian reconstruction approach, a pre-operative CT scan and a few known surrogate markers, we are able to accurately predict the position of the entire liver at all times. The surrogates may, for example, come from ultrasound, portal images captured during radiotherapy or from implanted electromagnetic beacons. In leave-one-out experiments, we achieve an average prediction error of 1.2 mm over sequences of 20 min with only three surrogates. Our model is accurate enough for clinically relevant treatment intervals and has the potential to be used for adapting the gating window in tumour therapy or even for tracking a tumour continuously during irradiation.

Keywords: respiratory motion, statistical model, prediction, tumour therapy, liver.

1 Introduction

Respiratory organ motion is a complicating factor in the treatment of liver tumours. Besides the superior/inferior (SI) motion caused by the diaphragm, there are secondary modes due to cardiac cycle motion, digestive activity, gravity and muscle relaxation, some of them causing a drift of the organ [10]. Non-rigid deformation during breathing introduces a significant amount of uncertainty in location during irradiation of a tumour. Studies have shown that 4-dimensional treatment planning is important for improved precision in radiotherapy [7], though not much work has been done in the field of precise liver motion estimation for non-invasive treatment of tumours. Rohlving *et al.* [5] acquired 3D-MRI liver data of exhalation, inhalation and eight time steps in-between. Deformation fields among them were obtained using non-rigid registration. These transformations were then applied to the vertex coordinates of geometrical models derived from the exhalation reference. More recently, He *et al.* [4] modeled 4D motion of lungs using Kernel PCA. They trained a support vector machine to model the relation between motion of fiducial markers on the lower abdomen/chest and the coefficients of the K-PCA. However, with the latter method, a correlation

between organ motion and the motion of an external surrogate may become unreliable over time in presence of organ drift [10]. Their approach is computationally involved and might therefore be inadequate for real-time reconstruction and tracking. Their reported average accuracy was 1.63 mm. Ehrhardt et al. [3] used diffeomorphic nonlinear intensity-based registration and the Log-Euclidean framework to build a motion model from thoracic 4D-CT lung data sets of 17 patients. The chosen approach is mathematically well formulated but it requires quite a few assumptions about breathing-depth and voxel intensities. Furthermore, the respiratory cycle is discretised to only four states and 4D-CT images of the patient are required for estimating a scaling factor for breathing depth. A prediction accuracy of 3.3 mm on average was achieved.

Most importantly, both previous and most other methods established inter-subject correspondence implicitly by co-registering all samples to a reference image, thereby relying on a correlation between shape and respiratory motion. Our approach, on the other hand, relies on mechanically corresponding points (*i.e.* points that move similarly), thus precisely focuses our analysis on the type of motion we wish to model. We capture this motion information by learning a PCA model of the deviation from exhalation position during quiet breathing. Thanks to the large amount of data used to learn our model and the applied prediction scheme, we achieve an average prediction error of only 1.2 mm.

This paper is organised as follows. In the next two sections we introduce the correspondence scheme and the statistical population model for respiratory organ motion, which is a generalisation of the statistical population-based model of intrafraction drift by von Siebenthal *et al.* [9]. We then describe the reconstruction approach used to predict the motion of a novel liver from only three known surrogates. Finally, our approach is evaluated using a series of leave-one-out experiments on 4D-MRI data.

2 Materials and Methods

2.1 Data Acquisition and Establishment of Correspondence

For this study, we used 4D-MRI sequences [8] of the liver from 12 healthy subjects (6 female, 6 male, average age 30, ranged 17-75). The in-plane resolution is 1.8 mm \times 1.8 mm and the slice thickness is 3-4 mm. The data was acquired over roughly one hour on 22 to 30 sagittal slices and a temporal resolution of 2.6-2.8 Hz. We extracted the deformation fields for each subject using the B-spline based non-rigid registration method proposed by Rueckert [6]. This process involved the manual segmentation of the liver in one master exhalation stack for each subject and results in dense deformation fields between this exhalation master and all respiratory states. In order to build a statistical model from this data, inter-subject correspondence had to be established. We used an approach similar to [11] that relies on mechanically corresponding points. Four landmark points were manually selected on each sagittal slice of the exhalation master, namely the delineations between the superior, anterior, posterior and inferior surfaces, see Fig. 1(a). The left liver lobe was kept out of the analysis because it is heavily

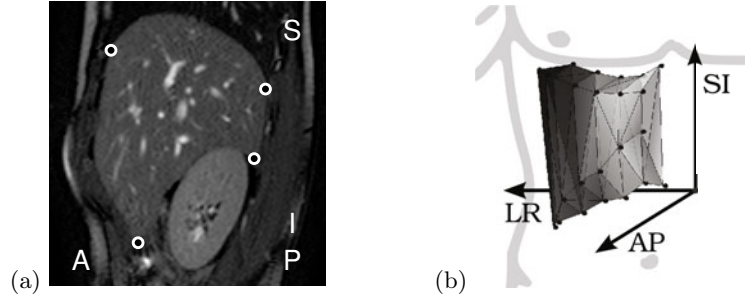


Fig. 1. (a) The location of a point relative to the liver surface is relevant for mechanical considerations. The white dots mark the delineations between the superior, anterior, posterior and inferior surfaces. (b) A coordinate system was chosen such that the LR-direction is perpendicular to the sagittal plane of the body and the AP-direction is aligned perpendicular to planes fitted to the anterior and posterior pairs of landmarks.

influenced by the motion of the heart, a factor that is not in the scope of this study. These points constituted a prototype shape for each subject, see Fig. 1(b). The prototypes were aligned inside the manually segmented master shape and the edges of the triangulation were projected onto the high resolution surface. That way, the master segmentations were remeshed and inter-subject correspondence of the surface was established. The total number of surface points after this step was 345 per liver. In order to establish correspondence for points within the liver, an isotropic grid with 10 mm resolution was placed in the average liver and then transformed to each of the remeshed surfaces using the Delaunay tetrahedrisation approach in [1], see Fig. 2. This finally gave a set of 12 topologically equivalent 3D liver volumes as well as dense spatio-temporal motion fields for each grid point.

2.2 Statistical Model

A liver instance is represented by a $3n$ -dimensional vector $\mathbf{v} = (x_1, y_1, z_1, \dots, x_n, y_n, z_n)$, where n corresponds to the number of vertices. To build a statistical model, 9 respiratory cycles were taken from each of the 12 subjects: 3 cycles from the beginning, middle and end of the acquisition session, respectively. The total number of volumes (12 subjects, 9 cycles, 8-20 steps per cycle) is $m=1312$. By distributing the selection of cycles all along the acquisition session, deformations of the liver due to organ drift over time are better represented. As our aim was to model organ motion and not organ shape, we removed the shape by taking the vector-field difference between each respiratory state and the subject's exhalation master state: $\mathbf{x} = \mathbf{v} - \check{\mathbf{v}}$. Principal Component Analysis (PCA) was then performed on the matrix of shape-neutralised offsets $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_m)$. This yields the orthonormal matrix of principal components $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_{m-1})$ and their corresponding eigenvalues $\lambda_1, \dots, \lambda_{m-1}$ that give the standard deviation σ_i of the principal components in descending order ($\sqrt{\lambda_i} = \sigma_i$).

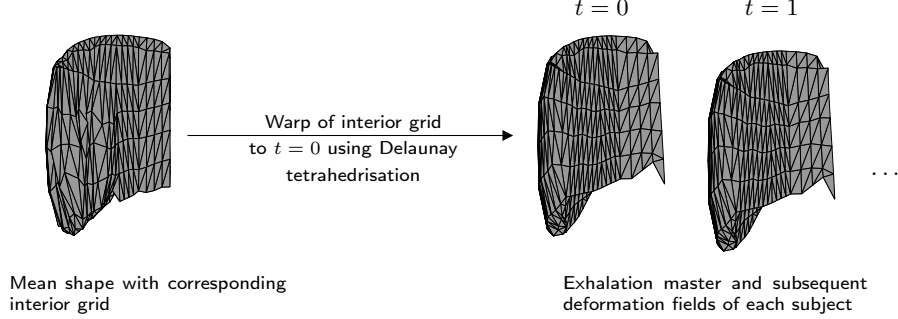


Fig. 2. An isotropic grid is positioned inside the mean shape and warped to the exhalation master shape $\tilde{\mathbf{v}}$ at $t = 0$ of each subject. From $t = 0$ to all forthcoming time steps, the deformation field is used to further warp the grid to any of the subsequent respiratory steps. This results in dense intra- and inter-subject correspondence.

We can transform (and thus decorrelate) the data by subtracting the mean offset vector $\mu = \frac{1}{m} \sum \mathbf{x}_i$, followed by a projection into model space:

$$\mathbf{c} = \text{diag}(\sigma_i^{-1}) \mathbf{U}^T (\mathbf{x} - \mu), \quad (1)$$

The deviation of a liver shape from its exhalation position during respiration can now be described in terms of our model,

$$\mathbf{v} = \tilde{\mathbf{v}} + \text{diag}(\sigma_i) \mathbf{U} \mathbf{c} + \mu. \quad (2)$$

From the observed partial information, we have to estimate a suitable \mathbf{c} that represents the motion information of the whole liver. In the next section, we will explain how this can be done.

2.3 Bayesian Motion Prediction from Sparse Measurements

We would like to use our model to predict the full motion of a liver based on only a few observed surrogate markers. In practice, these points may come, for example, from structures tracked in ultrasound images or from implanted electromagnetic beacons. We use the approach described in [2] to solve this problem. The partial observations are given by the vector $\mathbf{r} = \mathbf{L}(\mathbf{v} - \tilde{\mathbf{v}} - \mu)$, with a mapping $\mathbf{L} : \mathbb{R}^n \rightarrow \mathbb{R}^l, l < n$. Our aim is to find the model coefficient \mathbf{c} for the full vector \mathbf{x} that describes our partial measurement $\mathbf{r} = \mathbf{L}\mathbf{x}$. As we cannot expect to find an exact solution, we define the best reconstruction to be the one with minimal Mahalanobis distance $\|\mathbf{c}\|^2$, *i.e.* the one with highest prior probability. This can be formulated as a minimisation problem with regard to the model coefficients,

$$E = \|\mathbf{Q}\mathbf{c} - \mathbf{r}\|^2 + \eta \cdot \|\mathbf{c}\|^2, \quad (3)$$

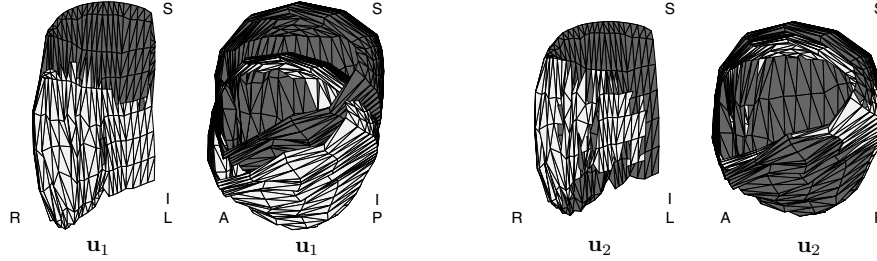


Fig. 3. Coronal and sagittal views of the mean liver shape deformed in direction of the first principal component \mathbf{u}_1 (left) and the second principal component \mathbf{u}_2 (right) of our respiratory model. The white and gray surfaces represent deformations of plus and minus $3\sigma_i$, respectively.

with $\mathbf{Q} = \mathbf{L}\mathbf{U} \cdot \text{diag}(\sigma_i)$. The regularisation factor η allows to trade off matching quality against prior probability. From the singular value decomposition $\mathbf{Q} = \overline{\mathbf{V}}\mathbf{W}\mathbf{V}^T$ we can calculate the most probable coefficients in a Bayesian sense,

$$\mathbf{c} = \mathbf{V} \text{diag}\left(\frac{w_i}{w_i^2 + \eta}\right) \overline{\mathbf{V}}^T \mathbf{r}. \quad (4)$$

The final shape can then be computed by projecting the model coefficient \mathbf{c} back into spatial domain according to Eq. (2). Note that \mathbf{Q} is of size $l \times \hat{m}$, with $\hat{m} \ll m$ the number of principal components used for reconstruction. Therefore, Eq. 4 can be easily solved in real-time.

3 Results

Figure 3 shows the mean and the variation of the first two principal components of our model. It can be seen that the first component clearly models SI motion. The additional deformation encoded by higher components quickly becomes very small. We further evaluated the quality and robustness of the model by measuring *expressiveness*, *compactness* and *generalisation*.

Compactness: Figure 4(a) shows a plot of the cumulative variance. The first 8 components already contain 98% of the total variability. It is desirable to have as much information as possible in the first modes, as this is a measure for the quality of the correspondence scheme. Also, a compact model suggests that the model is well suited to represent the specific class knowledge for which it was designed for.

Expressiveness: We computed leave-one-out models for all 12 subjects to evaluate the expressiveness of our model. For each model we projected the full shape of the unseen liver at inhalation into the model. This shows how well the model is able to describe (not predict) the most extreme respiratory state if it is entirely

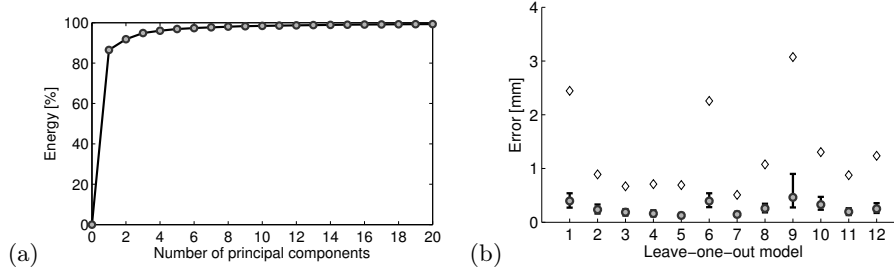


Fig. 4. (a) Cumulative energy plot of the statistical population model for the liver. (b) Median, 25th and 75th percentiles and maximum (diamond) of the projection error at inhalation.

known, and therefore gives an upper bound for model expressiveness, see Fig. 4(b). The lower and upper borders of the bar represent the 25th and the 75th error percentile, respectively, and the bar in between marks the median. The diamonds mark the maximum. Most grid points can be described with an error below 0.5 mm. On the other hand, some regions cannot be explained that accurately. However, as we look at the inhalation state, this is an upper bound for the projection error as the deviation from exhalation position is at its maximum.

Generalisation: We predicted the motion of 12 unseen livers in leave-one-out experiments. For all experiments, we assumed that a static exhalation shape of the left-out subject is given, *i.e.* a pre-operative CT scan. The prediction was based on only three surrogate markers for which we assume that we can track their deviation from exhalation position accurately. These points were selected at the inferior tip of Couinaud segment VI, at the diaphragm and near the center of the right liver lobe, in order to capture both the cyclic respiratory motion as well as organ drift that occurs mostly in the inferior part of the organ. We evaluated our method on a sequence of 20 min of motion data for each subject. A regularisation factor of $\eta = 2$ was chosen to avoid overfitting. As the error does not follow a Gaussian distribution, we give mean and maximum values as well as the 25th, 50th (median) and 95th percentile of the error evaluated at all 1326 grid points and over all respiratory states, see Tab. 1. The resulting average error over all experiments was 1.2 mm. For all subjects, the mean error was below 2 mm and for subjects 4, 5 and 7 it was significantly below 1 mm.

The used reconstruction method is a global optimisation and may not be optimal for single points, as can be seen in the 95th percentile that shows errors up to 5.6 mm. The maximum errors are outliers that originate from erroneous deformation fields in the ground truth data and had only a minor effect on the mean error. Furthermore, if all three surrogate markers are located in the proximity of the tumour, the prediction accuracy will be more stable in the region of the lesion, thus significantly lower maximum errors can be expected.

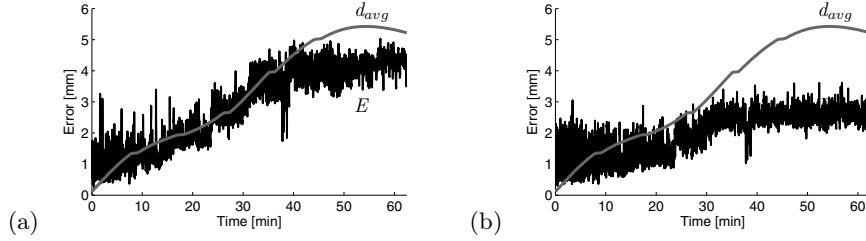


Fig. 5. Mean prediction error for a 60 min sequence (subject 12) that is subject to organ drift of up to $d_{avg} = 5.6$ mm (averaged over the entire liver). (a) Prediction with drift model that only contains states from the beginning of the acquisition session. (b) Prediction with model described in section 2.2.

Table 1. Motion prediction results obtained from leave-one-out experiments over 20 min of respiration and $\eta = 2$

	Prediction errors in [mm] per subject											
	1	2	3	4	5	6	7	8	9	10	11	12
Mean error	1.4	1.1	1.0	0.8	0.8	2.0	0.8	1.0	1.3	1.9	1.1	1.2
25th perc.	0.7	0.7	0.6	0.5	0.5	0.8	0.5	0.6	0.7	0.8	0.6	0.6
50th perc.	1.1	1.0	0.9	0.7	0.7	1.4	0.7	0.9	1.2	1.4	1.0	1.0
95th perc.	3.5	2.3	2.1	1.6	1.5	5.6	1.8	2.1	2.9	5.0	2.4	2.6
Max error	8.4	11.4	5.1	12.0	12.6	14.7	4.6	8.1	12.0	11.3	6.0	7.5

As some livers are subject to organ drift, their exhalation positions deform over time. In this case, they do not match the exhalation masters anymore which leads to higher prediction errors even at exhalation, see Fig. 5(a). However, since the model includes cycles from all over the acquisition session, most of the drift can be compensated, as Fig. 5(b) nicely shows.

4 Conclusions

We presented a generic statistical framework for the prediction of organ positions under respiratory motion from sparse measurements and demonstrated it for human livers. Our method accurately predicts the full respiratory motion over 20 min from very few observed points and also compensates for organ drift. The average error for each subject is between 0.8 mm and 2 mm and the average error over all experiments is 1.2 mm. With our population based approach, no time-consuming pre-computation (*e.g.* 3D image registration) is necessary to apply it to a novel patient, as it is the case for most subject-specific methods reported in the literature. Our method helps to significantly extend the gating window or even to irradiate a tumour continuously, allowing for more efficient and safer treatment. In forthcoming studies, we will increase the number of subjects to further generalise our model and enhance its expressiveness. Also, we

plan to relax the constraints on the surrogate markers, *e.g.* for tracking using implanted beads in portal images where one dimension is lost, or with ultrasound, where noisy measurements are expected. A further aspect in the clinical evaluation of our approach that has to be investigated, is the application to cancerous livers. The influence of a tumour on the elasticity and the deformation characteristics is not yet well understood.

Acknowledgement. This work was funded by the Swiss National Science Foundation (SNSF), project CRSII2_127549.

References

1. Barber, C.B., Dobkin, D.P., Huhdanpaa, H.: The quickhull algorithm for convex hulls. *ACM T. Math. Software* 22(4), 469–483 (1996)
2. Blanz, V., Vetter, T.: Reconstructing the complete 3D shape of faces from partial information. *Informationstechnik und Technische Informatik* 44(6), 295–302 (2002)
3. Ehrhardt, J., Werner, R., Schmidt-Richberg, A., Handels, H.: Statistical modeling of 4d respiratory lung motion using diffeomorphic image registration. *IEEE Transactions on Medical Imaging* 30(2), 251–265 (2011)
4. He, T., Xue, Z., Xie, W., Wong, S.T.C.: Online 4-D CT Estimation for Patient-Specific Respiratory Motion Based on Real-Time Breathing Signals. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) *MICCAI 2010 Part III*. LNCS, vol. 6363, pp. 392–399. Springer, Heidelberg (2010)
5. Rohlfing, T., Maurer Jr., C.R., O’Dell, W.G., Zhong, J.: Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated MR images. *Med. Phys.* 31(3), 427–432 (2004)
6. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: application to breast MR images. *IEEE T. Med. Imag.* 18(8), 712–721 (1999)
7. Shirato, H., Seppenwoolde, Y., Kitamura, K., Onimura, R., Shimizu, S.: Intrafractional tumor motion: lung and liver. *Semin. Radiat. Oncol.* 14(1), 10–18 (2004)
8. von Siebenthal, M., Cattin, P.C., Gamper, U., Lomax, A., Székely, G.: 4D MR Imaging Using Internal Respiratory Gating. In: Duncan, J.S., Gerig, G. (eds.) *MICCAI 2005*. LNCS, vol. 3750, pp. 336–343. Springer, Heidelberg (2005)
9. von Siebenthal, M., Székely, G., Lomax, A., Cattin, P.: Inter-Subject Modelling of Liver Deformation During Radiation Therapy. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part I*. LNCS, vol. 4791, pp. 659–666. Springer, Heidelberg (2007)
10. von Siebenthal, M., Székely, G., Lomax, A., Cattin, P.: Systematic errors in respiratory gating due to intrafraction deformations of the liver. *Med. Phys.* 34(9), 3620–3629 (2007)
11. Zsemlye, G.: Shape Prediction from Partial Information. Ph.D. thesis, ETH Zurich (2005)

Chapter 5

Robust Tumour Tracking From 2D Imaging Using a Population-Based Statistical Motion Model

In this study, presented at the *Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*, January 2012, Breckenridge, USA¹, various clinically-oriented simulations are performed using our population-based statistical motion model. Additionally, a new correspondence scheme is introduced. We study the case where 2d projections of fiducials, *e.g.* from x-ray images of implanted fiducial markers, undergo significant imaging noise and are used to continuously reconstruct organ motion. In particular, we study the effect of varying imaging angles on the reconstruction accuracy. The results show that an ideal angle for observing motion for model-based reconstruction is near-sagittal.

¹The article is available online at <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=06164749>.

Robust Tumour Tracking From 2D Imaging Using a Population-Based Statistical Motion Model

Frank Preiswerk, Patrik Arnold, Beat Fasel and Philippe C. Cattin
Medical Image Analysis Center
University of Basel, Switzerland

{frank.preiswerk,patrik.arnold,beat.fasel,philippe.cattin}@unibas.ch

Abstract

This paper describes a method for tracking a tumour using the planar projections of fiducial markers as surrogates. The projections can originate from various sources such as a beam-eye view X-ray, a portal imager or a fluoroscope. The two-dimensional position of the fiducial markers in the planar image in conjunction with a population-based statistical motion model is used to accurately predict and track the motion of a target volume during treatment. The basic assumption is that the projected surrogate locations contain valuable information about the in-plane motion of the lesion whereas the statistical motion model helps to describe the unobserved out-of-plane motion of the target volume. We analysed the accuracy with regard to varying the camera position and uncertainty in the measurement of the surrogate positions to simulate image noise and camera registration errors. The experiments showed that the tumour motion can be robustly predicted with an accuracy of 2.6 mm over a wide range of target volumes and treatment field directions despite a measurement error of $\sigma = 2$ mm for the fiducials.

1. Introduction

Respiratory organ motion is a complicating factor in the treatment of liver tumours. Besides the superior-inferior-motion (SI) caused by the diaphragm, there are secondary modes due to cardiac cycle motion, digestive activity, gravity and muscle relaxation, some of them causing a drift of the organ [13]. The deformation during breathing introduces a significant amount of uncertainty in location during irradiation of a tumour. The state-of-the art in tumour radiotherapy for motion management are breath-hold and respiratory gating. Both approaches are based on the observation that some phases of the breathing cycle, specifically exhalation and inhalation, are generally reproducible. Numerous publications have been written about respiratory organ

motion compensation [9, 13, 12, 4, 5, 3]. However, most of these approaches rely on idealised conditions like exact knowledge about the location of the fiducials, regular or stationary breathing patterns or they ignore the organ motion completely. Hostettler *et al.* [5] modeled abdominal viscera motion using volumetric meshes that slide inside an internal envelope. The envelope is deformed so that it matches the position of external skin markers. They achieve a prediction accuracy in the range of 2 to 3 mm. However, it has been shown that the correlation between external signals and organ motion is not always good, as it depends on the organ and even the particular region of the organ that is predicted [6]. For the liver we can expect a high correlation near the diaphragm but a lower correlation in the inferior region. Additional perturbation comes from the heartbeat. Arnold *et al.* [1] use a motion atlas in conjunction with a statistical drift model to predict liver motion in High Intensity Focused Ultrasound (HIFU) scenarios. They temporally predict the liver position 300 ms into the future. While the motion atlas temporally predicts the intra-cycle respiratory motion, the statistical model compensates for the organ drift over longer time scales.

We have shown previously that a statistical motion model can not only describe organ drift but also the complete 3D organ motion based on a small number of known surrogate positions [7]. In this work, we apply a statistical model to infer the complete 3D motion - in particular the out-of-plane motion - from tracked 2D marker positions. This study is one step closer towards a practical application of model-based respiratory motion compensation, as much of the prior work in this field was performed under idealised conditions, where the 3D positions of surrogates are assumed to be known. In this paper we examine some of the practical issues in motion compensation. In particular, we examine the effect of 3D to 2D projection of the observed fiducials as it happens with most intra-operative imaging techniques for fiducial observation. Furthermore we study the effect of imprecise localisation of the said fiducials because of image noise and/or imperfect camera cali-

bration. Having a focus on applicability, we additionally introduce a novel correspondence scheme (as compared to the method described in [12]) that significantly speeds up the segmentation and correspondence establishment step when applying the statistical model to a novel patient. We extensively evaluate our method in a series of cross-validation experiments.

2. Materials and Methods

In this section we will first describe the statistical motion model and our new correspondence establishment method. Then we will explain how 3D positions can be inferred from the tracked 2D markers.

2.1. Data and Correspondence

For this study, we used 4D-MRI sequences [11] of the liver from 20 healthy subjects (mixed genders, average age 30, range 17-75). The in-plane resolution is $1.8 \text{ mm} \times 1.8 \text{ mm}$ and the slice thickness is 3 to 4 mm. The data was acquired over roughly one hour on 22 to 30 sagittal slices with a temporal resolution of 2.6 to 2.8 Hz. We extracted the deformation fields for each subject using the B-spline based non-rigid registration method proposed by Rueckert [10]. This process involved the manual segmentation of the liver in one master exhalation stack for each subject and results in dense deformation fields between this exhalation master and all respiratory states. For this study we used a similar statistical population model as described in [13] and [7]. However, we developed an improved correspondence scheme that makes it easy to apply the model to a new patient. In a preparatory step, the liver of each patient was manually segmented on each sagittal slice of a pre-operative planning CT or MRI scan. Additionally, four points were manually selected on each slice to define the correspondence: Two points where the liver is attached to the anterior and posterior wall, respectively, as well as the highest point on the superior end (diaphragm) and the lowest point at the inferior end of the organ, see Fig. 1 (left). The reason for choosing the AP points is because the discontinuity between fixed tissue and the moving organ impacts the way the organ can move. The SI points were selected because the main mode of motion is in this direction. We applied an arc-length discretisation algorithm that combines mask and landmarks to resample the organ contour for each slice. This way, four individual segments with equidistant points are generated, see Fig. 1 (middle, right). The number of points is controlled individually for each segment. We chose 5 points for both superior segments and 10 points for both inferior segments. For each of the 30 segment points we define a smoothing spline controlled by the corresponding point on all slices, $P_i = (p_{i,1}, p_{i,2}, \dots, p_{i,k})$, that is, in out-of-plane (or left-right) direction, where k is the total number of slices. The spline is then resampled at 30 loca-

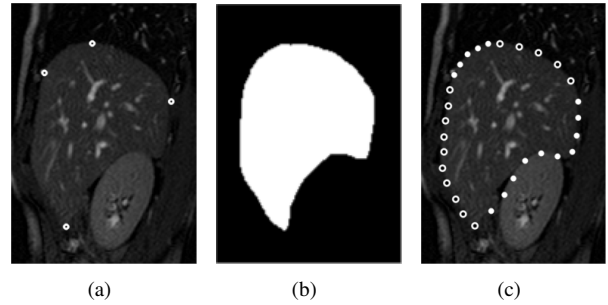


Figure 1. (a) Manually selected landmarks for correspondence on a selected sagittal slice. (b) Manually defined mask. (c) Automatically generated resampling of the liver contour based on mask and landmarks. The alternating markers highlight the four individually controlled segments.

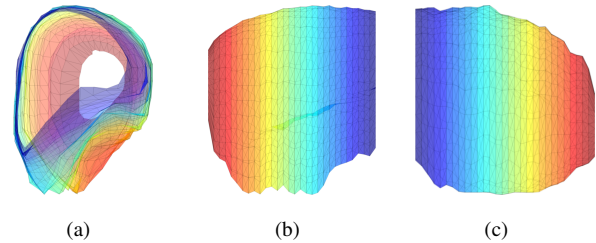


Figure 2. (a) sagittal, (b) coronal and (c) axial view of a final mesh after resampling and interpolation.

tions. This is done because image acquisitions of different subjects can consist of different numbers of slices n . After resampling the splines, the approach is independent of n .

2.2. Setup

Our setup is as follows: We assume that the imaging device and the patient table are registered, *i.e.* a coordinate transform \mathbf{A}_c consisting of rotation and translation between the systems is known. Additionally, we assume that the fiducials of interest can be tracked in the sequence of images, *i.e.* their location in the planar projections can be determined with a certain accuracy. Lastly, we assume that a pre-operative 3D image (CT or MR) of the patient’s liver is available that can be used to perform the correspondence establishment steps described in Sec. 2.1 to align the patient with our model. This yields the transform \mathbf{A}_m between model and patient. The latter is not a restricting condition, as 3D CT datasets are routinely acquired for treatment planning anyways. Figure 3 depicts the involved devices and coordinate systems.

An interesting case that is worth mentioning is a beam-eye view camera that allows to observe the target from the same viewpoint and angle as the irradiation device, thus minimising the camera/irradiation device calibration error. Without restricting generality we assumed an orthogonal

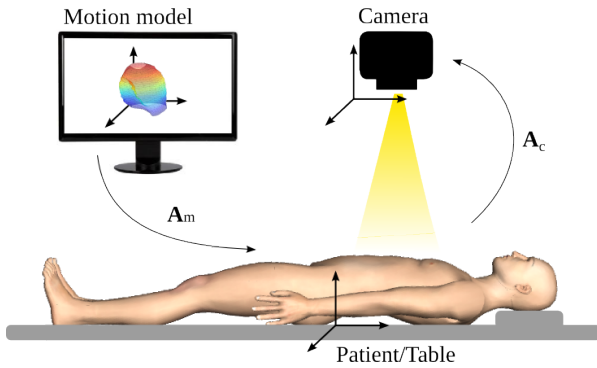


Figure 3. Overview of the three coordinate systems involved in our simulations. We assume that the patient table and the imaging device are registered and thus the camera matrix is known. Additionally, we assume that the transformation of the motion model to the patient liver is obtained. In a real-world scenario this can be done through a pre-operative planning CT image via the patient table coordinate system.

projection model in this research. A point $p = (x, y, z)$ in the patient's liver can be transformed to the image plane by multiplying it with the camera matrix $q = \mathbf{A}_c p$ and subsequently removing the z -dimension. We have generated the input data for our experiments by projecting sequences of model gridpoints (from Sec. 2.1) to camera space that way for each subject. The result are temporally resolved 2D fiducial positions. This is equivalent to the case where implanted fiducial markers are tracked in planar images during treatment. Furthermore, we added Gaussian noise to the projected locations to simulate all types of tracking error sources.

2.3. Statistical Model and Reconstruction

In this section we shortly summarise the most important steps when building a statistical motion model and using it for inferring (reconstructing) tumour positions.

In our statistical model, a liver instance is represented by a $3n$ -dimensional vector $\mathbf{v} = (x_1, y_1, z_1, \dots, x_n, y_n, z_n)$, where n corresponds to the number of model vertices. After removing the shape information by taking the vector-field difference between each respiratory state \mathbf{v} and the subject's exhalation master state $\tilde{\mathbf{v}}$ (from Sec. 2.1), we get the motion data that is used to build the model: $\mathbf{x} = \mathbf{v} - \tilde{\mathbf{v}}$. Principal Component Analysis (PCA) on all samples $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_m)$ yields the orthonormal matrix of principal components $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_{m-1})$ and their corresponding Eigenvalues $\lambda_1, \dots, \lambda_{m-1}$ that represent the standard deviation σ_i of the principal components in descending order ($\sqrt{\lambda_i} = \sigma_i$). The basic idea behind our approach is that the observed fiducial positions $q_i = (x_i, y_i)$ also contain a significant amount of information regarding the depth variation (z -dimension) that is lost when capturing the 2D images. Our statistical model captures these correlations

and allows to recover the most plausible depth-change in a Bayesian sense (see Fig. 5). As described in [2], we can compute a reconstruction of the model coefficient vector \mathbf{c} , *i.e.* the complete motion field, from any number of measured 2D points p_i using the following equation:

$$\mathbf{c} = \mathbf{V} \text{diag}\left(\frac{w_i}{w_i^2 + \eta}\right) \bar{\mathbf{V}}^T \mathbf{r} \quad (1)$$

with

$$\mathbf{A}_c \mathbf{U} \cdot \text{diag}(\sigma_i) = \bar{\mathbf{W}} \mathbf{W} \mathbf{V}^T, \quad (2)$$

where $r = (q_{1,x}, q_{1,y}, q_{2,x}, q_{2,y}, \dots, q_{n,x}, q_{n,y})'$ is the vector of partial information and w_i being the elements of the diagonal matrix \mathbf{W} . The deviation of a liver shape from its exhalation position during respiration can now be described in terms of our model:

$$\mathbf{v} = \tilde{\mathbf{v}} + \text{diag}(\sigma_i) \mathbf{U} \mathbf{c} + \mu. \quad (3)$$

where μ is the mean motion vector $\frac{1}{m} \sum \mathbf{x}_i$.

2.4. Experiments

For each of our 20 subjects, we assessed how well our approach can describe the 3D motion of the target volume for 5 min duration, based solely on a sequence of 2D fiducial locations (and the pre-operative CT scan). We generated the 2D projections of three fiducials for all experiments as described in Sec. 2.2. Figure 4 depicts the location of these markers on the mean shape. Naturally, the reconstruction quality increases with the number of known points. However, in practice the number of observable points is limited and three fiducials have shown to be a reasonable number. Compared to the work in [7] where motion fields are reconstructed from 3D points to prove that statistical models are appropriate for motion prediction, we focus on a more practical application and also use a more generalised statistical motion model comprising of 20 subjects instead of just 12. Additionally, we selected 30 instead of 3 cycles from the beginning, middle and end of the acquisition session, respectively, as well as the 30 deepest cycles per subject, totalling 120 cycles per subject.

We generated a sequence of 2D projections of the fiducials for each patient and evaluated our approach using leave-one-out models. The camera position in the first experiment was chosen to acquire sagittal images of the observed fiducials, as it has been shown that the LR motion is smaller than AP motion for the liver [8]. This way, the extent of out-of-plane motion is minimised. The duration of each sequence was 5 min. We predicted the motion of all model grid points within a radius of $r = 15$ mm around the centroid of the three observed fiducials, see Fig. 5. We also studied the effect of varying this radius. In another set of experiments we were interested in the question of which

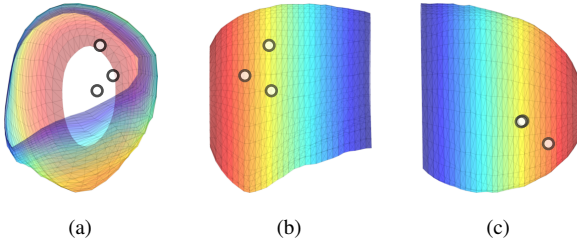


Figure 4. (a) sagittal, (b) coronal and (c) axial view of the mean shape with the selected fiducials highlighted.

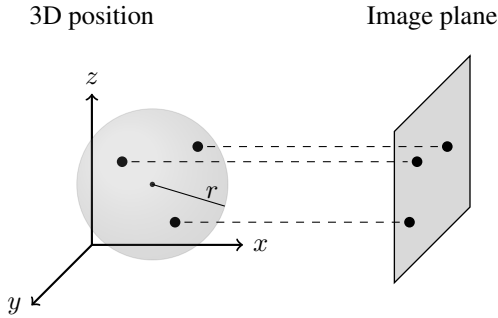


Figure 5. The depth information of the fiducials is lost in a 2D projection. The input for our algorithm are points on the 2D image plane. Our method reconstructs the 3D motion of the observed points as well as all other grid points within a sphere of radius $r = 15$ mm around the centroid of the three fiducials.

viewing angle is best for 2D imaging. As previously mentioned, LR motion is small for livers [8], so a natural position for the camera is sideways from left or right toward the patient. We varied the camera position around the patient as depicted in Fig. 10 to assess the results of different viewpoints.

In a real scenario, there are registration errors between the patient, table and the imaging device. Therefore, the matrix \mathbf{A} will never be entirely correct. Additionally, the points can only be tracked up to a certain accuracy, due to image tracking errors and resolution constraints. Because of these sources of error, we added Gaussian noise with a standard deviation of $\sigma = 2$ mm to the tracked 2D locations. We compare all our results to a rigid shift of the target volume, governed by the centroid of the tracked fiducials. This is important in order to study how much additional precision the model yields over a naive rigid tracking scheme.

3. Results

To better study the benefits of our proposed motion management approach, we established a silver-standard for comparison. In this simplified prediction scheme we assume a rigid motion model without any out-of-plane components. To compute the offset vector s , the displacement of the fiducial's centre of mass was used. Let m be this

centre of mass and m_0 the same centre of mass in the exhalation master. The offset $s = m - m_0$ was added to all points in the target volume of the exhalation master to predict their position in each breathing state. Figure 7 depicts the resulting error for this prediction scheme. The mean error over all subjects is 3.7 mm and the maximum is 13 mm (Subject 12).

In the first experiment, we predicted the motion of the target volume with our approach from a simulated camera that acquires sagittal images and compare it to the silver-standard. The results are depicted in Fig. 8. The mean error over all subjects was 2.6 mm and the maximum was 12 mm (Subject 19). In comparison with the reference results above, this shows that the prediction based on our method improves the accuracy by 1.1 mm on average over the silver-standard. This significant improvement is quite surprising as the sagittal projection angle is expected to yield the smallest prediction error for the rigid assumption, since the out-of-plane motion is the smallest in this viewing plane. The maximum errors are mainly caused by a relatively small number of outliers from abnormal breathing cycles.

A major advantage of our method is its ability to deal with errors in measurement. The Gaussian noise that is added to the location of the fiducials heavily influences the results of a rigid prediction scheme. To point this out, we repeated the rigid prediction experiments without adding noise. The resulting mean error dropped to 1 mm (compared to 3.7 mm with $\sigma = 2$ mm). Thus, a rigid prediction scheme only yields a very small error under idealised conditions. However, this is never the case in real applications. Uncertainties caused by image noise and imperfect camera calibration need to be taken into account. Our statistical model can deal with this kind of uncertainty as it always gives a plausible reconstruction of the motion field.

To see how the error evolves with increasing size of the predicted volume, we varied the radius r from 10 mm to 40 mm. Figure 6 depicts the results for both model-based prediction as well as the rigid prediction scheme. It shows that we can robustly predict even large target volumes without compromising the prediction accuracy. Although the silver-standard also shows no significant dependence on the radius, it performs consistently worse by more than 1 mm.

In a final experiment, we performed an exhaustive evaluation of all combinations of azimuth $az \in [-90, 90]$ and elevation $el \in [-90, 90]$ in steps of 5° with $el = az = 0^\circ$ being the sagittal view. This resulted in a total of 1369 separate experiments per subject with different camera positions. Figure 9 shows the results. The prediction error is lowest for those perspectives that allow to observe the SI motion of the liver. Perspectives that point in a similar direction as this main mode of variation tend to give higher prediction errors. This is because the SI motion nearly van-

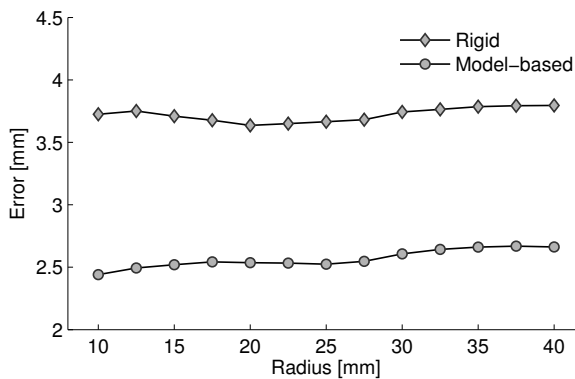


Figure 6. Mean prediction error as a function of the target volume radius. Both errors increase only very little, however, the model-based prediction is significantly more precise for any size.

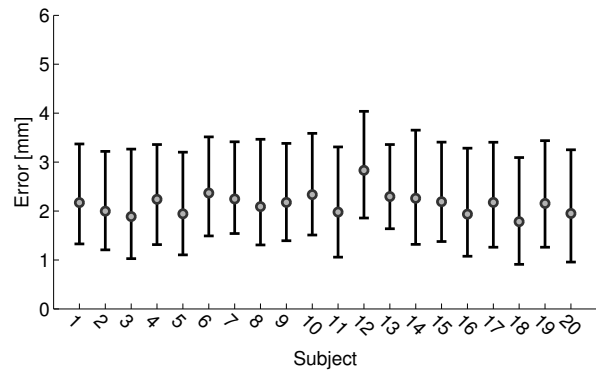


Figure 8. Results with sagittal imaging. Errorbars represent the median as well as the 25th and 75th percentiles. Compared to the results in Fig. 7, our method significantly reduces the prediction error for all subjects.

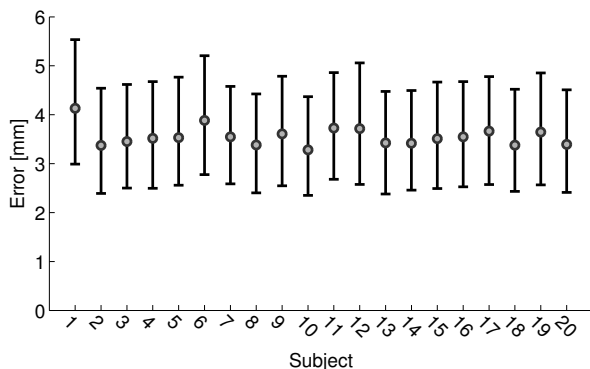


Figure 7. Precision of rigid motion compensation with sagittal imaging (silver standard). Errorbars represent the median as well as the 25th and 75th percentiles.

ishes when observed from these angles. The precision is largely insensitive to variation of the azimuth, as the SI motion can be observed from all around the patient’s body. The position of the peaks in the error surface in Fig. 9 correspond to the direction of the motion fields in our data. Figure 11 depicts an intensity map of the same surface for Subject 1 compared to the mean motion field of Subject 1. This again clearly points out how a view perpendicular to the motion yields a better reconstruction and, conversely, how a camera position parallel to the main direction of motion makes the observable variation almost disappear and therefore yields bad reconstructions, *i.e.* motion estimations.

4. Conclusion

We have described an application of a statistical motion model to infer the 3D motion of a target volume from surrogate markers projected onto planar images and evaluated the method on a realistic treatment scenario. We used a population-based statistical motion model of 20 subjects and simulated a series of 2D fiducial marker images. Gaussian noise was added to the 2D positions to take into ac-

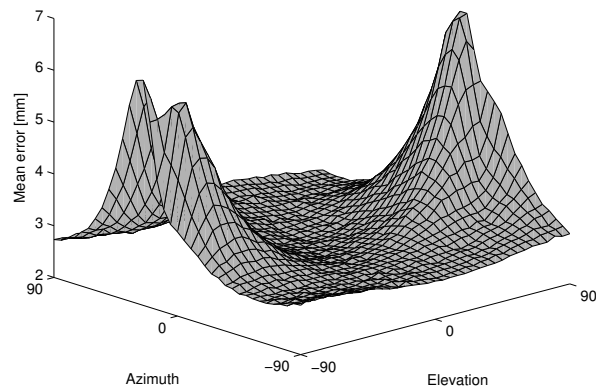


Figure 9. This surface shows the mean error over all subjects as a function of the camera position. It can be nicely seen that the prediction accuracy drops significantly when *el* approaches the values $+90^\circ$ and -90° .

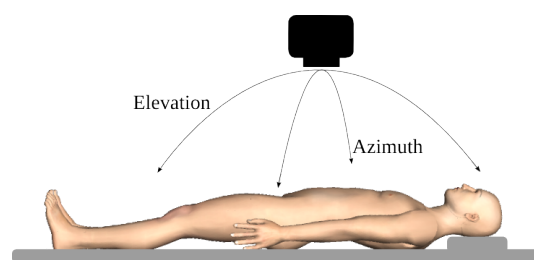


Figure 10. We evaluated a large number of camera positions to study its importance for motion prediction.

count uncertainties in tracking that arise in real-world scenarios, such as calibration errors or image noise. From these 2D input points we predicted the motion of target volumes of 10 to 40 mm radius in sequences of 5 min of respiratory motion for each subject. We compared our approach to a rigid prediction scheme and could demonstrate that our

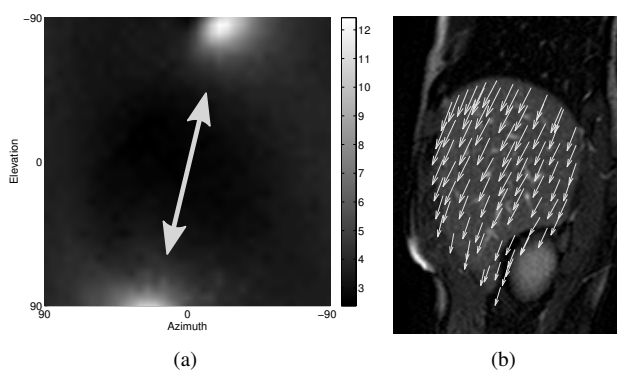


Figure 11. (a) The prediction error map of Subject 1 from the evaluation of all camera positions compared to the mean motion of Subject 1 on a sagittal slice (b). This shows how the main direction of respiratory motion influences the prediction accuracy from different viewpoints. Good camera viewpoints are approximately orthogonal to the main direction of motion.

method performed significantly better in predicting the motion of the target volume. In a second experiment, we evaluated a large number of camera positions to study the effect of camera angles on the motion prediction accuracy. As expected, it turned out that the best camera positions are perpendicular to the main mode of motion. When fiducials are observed in a sagittal view, the prediction results are best. The results became dramatically worse when approaching a near-axial view.

Tumour tracking for cancer treatment is a field that needs precise and careful methods and thorough evaluations. The simulations performed in this research were designed to be similar to a real application scenario, in order to evaluate some of the practical aspects of using a statistical motion model for motion compensation. We expect that our model performs well on pathological cases too, because tumours tend to reduce the tissue flexibility and extent of motion, thereby reducing the possibility of observing large deformations that may lie outside the span of our model. However, this assumption should be validated in the future. Non-invasiveness is at the centre of our research and we believe that our method can help to simplify tumour treatment and minimise collateral damage even in the absence of implanted fiducial markers. Therefore we are currently investigating the possibility of replacing the implanted fiducial landmarks by natural landmarks tracked by ultrasound, since our goal is to rely solely on non-invasive imaging techniques.

References

- [1] 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 *MICCAI*, volume LNCS 6892, pages 623–630, 2011. 209
- [2] V. Blanz and T. Vetter. Reconstructing the Complete 3D Shape of Faces from Partial Information. *Informationstechnik und Technische Informatik*, 44(6):295–302, 2002. 211
- [3] J. Ehrhardt, R. Werner, A. Schmidt-Richberg, and H. Handels. Statistical Modeling of 4D Respiratory Lung Motion Using Diffeomorphic Image Registration. *Medical Imaging, IEEE Transactions on*, 30(2):251–265, feb. 2011. 209
- [4] T. He, Z. Xue, W. Xie, and S. T. C. Wong. Online 4-D CT Estimation for Patient-Specific Respiratory Motion Based on Real-Time Breathing Signals. In *Proc. MICCAI 2010*, pages 392–399. Springer, 2010. 209
- [5] A. Hostettler, S. A. Nicolau, Y. Rmond, J. Marescaux, and L. Soler. A real-time predictive simulation of abdominal viscera positions during quiet free breathing. 209
- [6] M. J. Murphy. Tracking Moving Rrgans in Real Time. *Semin. Radiat. Oncol.*, 14(1):91–100, 2004. 209
- [7] F. Preiswerk, P. Arnold, B. Fasel, and P. C. Cattin. A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements. *Abdominal Imaging 2011*, Springer LNCS 7029, 2011. 209, 210, 211
- [8] T. Rohlfing, C. R. Maurer, Jr., W. G. O’Dell, and J. Zhong. Modeling Liver Motion and Deformation During the Respiratory Cycle Using Intensity-Based Free-Form Registration of Gated MR Images. *Proc. SPIE Med. Imag. 2001*, 4319(1):337–348, 2001. 211, 212
- [9] T. Rohlfing, C. R. Maurer, Jr., W. G. O’Dell, and J. Zhong. Modeling Liver Motion and Deformation During the Respiratory Cycle Using Intensity-Based Nonrigid Registration of Gated MR Images. *Med. Phys.*, 31(3):427–432, 2004. 209
- [10] 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. Imag.*, 18(8):712–721, 1999. 210
- [11] 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, 2007. 210
- [12] M. von Siebenthal, G. Székely, A. Lomax, and P. Cattin. Inter-Subject Modelling of Liver Deformation During Radiation Therapy. In *MICCAI 2007*, volume 4791 of LNCS, pages 659–666. Springer, 2007. 209, 210
- [13] M. von Siebenthal, G. Székely, A. Lomax, and P. Cattin. Systematic Errors in Respiratory Gating due to Intrafraction Deformations of the Liver. *Med. Phys.*, 34(9):3620–3629, 2007. 209, 210

Chapter 6

Towards More Precise, Minimally-Invasive Tumour Treatment Under Free Breathing

This invited publication was presented at the 34th *International Conference of the IEEE Engineering in Medicine and Biology Society* (EMBC), August 2012, San Diego, USA¹. In this work we compare the prediction performance using different kinds of surrogate signals. We compare the use of simulated 3d, 2d and 1d surrogate signals for motion prediction using our motion model and show that even in the 1d case, the accuracy is still within a clinically acceptable range.

¹The article is available online at <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=06346782>.

Towards More Precise, Minimally-Invasive Tumour Treatment Under Free Breathing

Frank Preiswerk, Patrik Arnold, Beat Fasel and Philippe C. Cattin¹

Abstract—In recent years, significant advances have been made towards compensating respiratory organ motion for the treatment of tumours, *e.g.* for the liver. Among the most promising approaches are statistical population models of organ motion.

In this paper we give an overview on our work in the field. We explain how 4D motion data can be acquired, how these motion models can then be built and applied in realistic scenarios. The application of the motion models is first shown on a case where 3D surrogate marker data is available. Then we will evaluate the prediction accuracy if only 2D and lastly 1D surrogate marker motion data is available. For all three scenarios we will give quantitative prediction accuracy results.

I. INTRODUCTION

Respiratory organ motion is a complicating factor in the treatment of liver tumours. Non-rigid deformation during breathing introduces a significant amount of uncertainty in location during irradiation of a tumour. It has been shown that 4D treatment planning is important for improved precision in radiotherapy [10]. A lot of research has been done for handling respiratory motion of organs. Some of the approaches rely on a correlation between external markers and internal organ motion [5], which is not always valid either due to organ drift [12], [6] or because of varying motion patterns at different positions of the organ. The approach by Ehrhardt *et al.* [4] is mathematically well formulated but requires quite a few assumptions about breathing-depth and voxel intensities and falls short in the temporal resolution of the model. Our approach to this problem is based on the acquisition of Magnetic Resonance Imaging (MRI) sequences of volunteers and a number of algorithms for processing this data. First, we retrospectively reconstruct 4D images in order to generate a 4D-MRI sequence for each subject. Then we register the data to obtain deformation fields for each sequence and establish mechanical correspondence among all subjects. We finally capture the motion information by learning a PCA model of the deviation from exhalation position during quiet breathing. Finally, we apply a prediction algorithm to sparse data in order to obtain a reconstruction of the motion for the whole organ at any point in time. Our model can equally well handle 1D, 2D or 3D input information. Even if only a single 1D signal (*i.e.* the Superior/Inferior (SI) motion component of a single point on the diaphragm) is available, our model provides an accurate reconstruction of the position of the complete organ.

In this paper we shortly wrap up the basics of our modeling pipeline. We provide quantitative results of two

of our most recent studies and additionally give the latest results on using our model in a gated setting with only a 1D surrogate.

II. STATISTICAL MOTION MODEL

Building a statistical motion model comprises of a number of steps. First, the image data is obtained using an MRI scanner and the data is registered to obtain deformation fields over time. Then, correspondence is established among all subjects. Finally, the deformation fields are processed using Principal Component Analysis (PCA) to compute the model.

A. Data Acquisition

We used 4D-MRI sequences [11] of the liver from 20 healthy subjects. The in-plane resolution is in the range of 1.5 mm and the slice thickness is 3 – 4 mm, depending on the subject. The data was acquired over roughly one hour on 22 to 30 sagittal slices and a temporal resolution of 2.6 – 2.8 Hz. 4D-MRI sequences are generated by retrospectively stacking the acquired 2D image slices based on the navigator technique as described in [11], and depicted in Fig. 1.

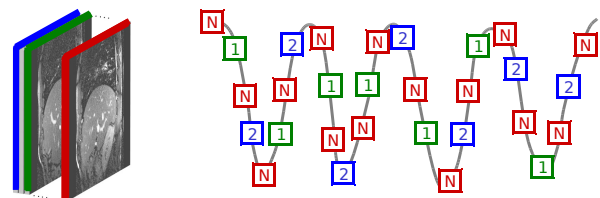


Fig. 1. The 4D-MRI acquisition scheme. Each slice position (1, 2, ...) is temporally encapsulated by two navigators (N). The navigator slice is always acquired at the same spatial position and defines a phase for each encapsulated image. This allows to find matching slices of the same breathing phase for each position and thus gives a 3D stack for every point in time.

The 3D deformation fields are then extracted using the B-spline based non-rigid registration method proposed by Rueckert *et al.* [9]. This process involved the manual segmentation of the liver in one master exhalation stack for each subject and results in dense deformation fields between this exhalation master and all respiratory states, see Fig. 2.

B. Establishment of Correspondence

In order to build a statistical model from the deformation fields, inter-subject correspondence has to be established. We developed a correspondence scheme that makes it easy to apply the model to a new subject [8]. In a preparatory step, the liver of each subject was manually segmented on each

¹Medical Image Analysis Center, University of Basel, Switzerland
frank.preiswerk at unibas.ch

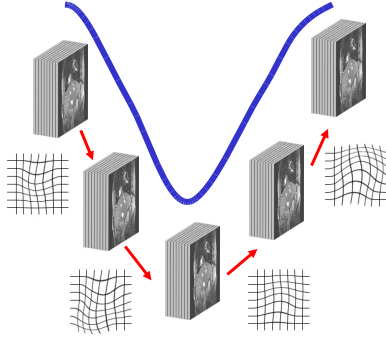


Fig. 2. Non-rigid registration of the 4D-MRI stacks results in dense spatio-temporal motionfields.

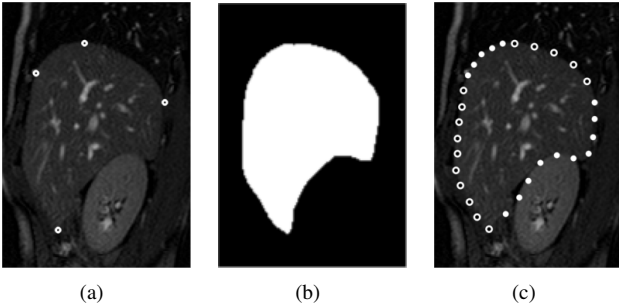


Fig. 3. (a) Manually selected landmarks for correspondence on a selected sagittal slice. (b) Manually defined mask. (c) Automatically generated resampling of the liver contour based on mask and landmarks. The alternating markers highlight the four individually controlled segments.

sagittal slice of one MRI scan. Additionally, four points were manually selected on each slice to define the correspondence: Two points where the liver is attached to the anterior and posterior wall, respectively, as well as the highest point on the superior end (diaphragm) and the lowest point at the inferior end of the organ, see Fig. 3(a). We applied an arc-length discretisation algorithm that combines the mask and landmarks to resample the organ contour for each slice, see Fig. 3(b)&3(c). The data is then again resampled in z -direction because image acquisitions of different subjects can consist of different numbers of slices. In order to establish correspondence for points within the liver, an isotropic grid was placed in the average liver and then transformed to each of the remeshed surfaces using the Delaunay tetrahedrisation approach [2], see Fig. 4. This finally gave a set of 20 topologically equivalent 3D liver volumes that constitute one exhalation master $\tilde{\mathbf{v}}$ for each subject and combined with the registration results, dense spatio-temporal motion fields for each grid point.

C. Statistical Model and Reconstruction

In our statistical model, a liver instance is represented by a $3n$ -dimensional vector $\mathbf{v} = (x_1, y_1, z_1, \dots, x_n, y_n, z_n)$, where n corresponds to the number of model vertices. After removing the shape information by taking the vector-field difference between each respiratory state \mathbf{v} and the subject's exhalation master state $\tilde{\mathbf{v}}$, we get the motion data that is used

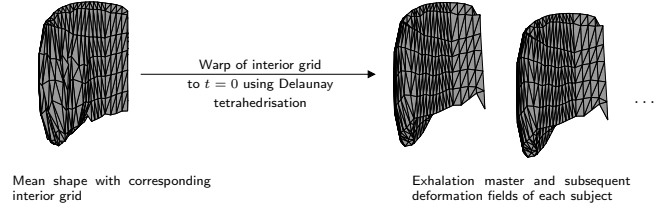


Fig. 4. An isotropic grid is positioned inside the mean shape and warped to the exhalation master shape $\tilde{\mathbf{v}}$ at $t = 0$ of each subject. From $t = 0$ to all forthcoming time steps, the deformation field is used to further warp the grid to any of the subsequent respiratory steps. This results in dense intra- and inter-subject correspondence.

to build the model: $\mathbf{x} = \mathbf{v} - \tilde{\mathbf{v}}$. Principal Component Analysis (PCA) on all samples $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_m)$ yields the orthonormal matrix of principal components $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_{m-1})$ and their corresponding Eigenvalues $\lambda_1, \dots, \lambda_{m-1}$ that represent the standard deviation σ_i of the principal components in descending order ($\sqrt{\lambda_i} = \sigma_i$). We can transform (and thus decorrelate) the data by subtracting the mean offset vector $\mu = \frac{1}{m} \sum \mathbf{x}_i$, followed by a projection into model space:

$$\mathbf{c} = \text{diag}(\sigma_i^{-1}) \mathbf{U}^T (\mathbf{x} - \mu), \quad (1)$$

The deviation of a liver shape from its exhalation position during respiration can now be described in terms of our model,

$$\mathbf{v} = \tilde{\mathbf{v}} + \text{diag}(\sigma_i) \mathbf{U} \mathbf{c} + \mu. \quad (2)$$

From the observed partial information, we have to estimate a suitable model coefficient vector \mathbf{c} that represents the motion information of the whole liver. In practice, this information may come, for example, from structures tracked in ultrasound, x-ray or portal images or from implanted electromagnetic beacons. We use the approach described in [3] to solve this problem. The partial observations are given by the vector $\mathbf{r} = \mathbf{L}(\mathbf{x} - \mu)$, with a mapping $\mathbf{L} : \mathbb{R}^n \rightarrow \mathbb{R}^l, l < n$. Our aim is to find the model coefficient \mathbf{c} for the full vector \mathbf{x} that describes our partial measurement. As we cannot expect to find an exact solution, we define the best reconstruction to be the one with minimal Mahalanobis distance $\|\mathbf{c}\|^2$, *i.e.* the one with highest prior probability. This can be formulated as a minimisation problem with regard to the model coefficients,

$$E = \|\mathbf{Q}\mathbf{c} - \mathbf{r}\|^2 + \eta \cdot \|\mathbf{c}\|^2, \quad (3)$$

with $\mathbf{Q} = \mathbf{L}\mathbf{U} \cdot \text{diag}(\sigma_i)$. The regularisation factor η allows to trade off matching quality against prior probability. From the singular value decomposition $\mathbf{Q} = \bar{\mathbf{V}}\mathbf{W}\mathbf{V}^T$ we can calculate the most probable coefficients in a Bayesian sense,

$$\mathbf{c} = \mathbf{V} \text{diag}\left(\frac{w_i}{w_i^2 + \eta}\right) \bar{\mathbf{V}}^T \mathbf{r}. \quad (4)$$

The final shape can then be computed by projecting the model coefficient \mathbf{c} back into spatial domain according to Eq. (2). Note that \mathbf{Q} is of size $l \times \hat{m}$, with $\hat{m} \ll m$ the number of principal components used for reconstruction. Therefore, Eq. (4) can be easily solved in real-time.

III. EXPERIMENTS

Given the statistical motion model, it can be applied for predicting the tumour location from surrogate markers. In the following we first evaluate our model when a set of 3D surrogate markers is available. Then we progressively increase the difficulty by predicting only based on 2D surrogate marker information and lastly based on a simple 1D signal.

A. Prediction from Sparse 3D Data

In the first scenario [7] we assume that surrogate markers can be tracked in three dimensions. This scenario is applicable to a treatment *e.g.* with the Calypso System (Calypso Medical Technologies, Inc) which provides accurate real-time 3D positions of implanted markers. Similar data can also be produced by bi-plane X-rays such as in the CyberKnife Robotic Radiosurgery System (Accuray, Inc.). We obtained an average error of 1.2 mm. Furthermore, the 95th percentile of the errors was at 2.8 mm and the large errors appeared mainly at inhalation. This suggests that the error could be additionally reduced using amplitude gating, where no prediction (and thus no treatment) is performed near full inhalation, see Fig. 8

B. Prediction from Sparse 2D Data under Projection and Noise

In the next scenario we evaluated the motion models in a set-up where only 2D motion information is available [8]. Clinically, this applies to cases where ultrasound or portal images are used to observe the organ during treatment. We simulated such 2D data by projecting the 3D positions to 2D planes under a large number of different projection angles and additionally superimposed Gaussian noise ($\sigma = 2$ mm) in order to simulate for various sources of errors that would arise in a real-world setup. The model is able to accurately predict the overall 3D deformation field, similarly to the case where 3D surrogate information is available. Our method is sufficiently robust to produce accurate reconstructions even in presence of noise. For the sagittal view, we obtained an average error of 2.6 mm. Figure 5 clearly shows that this error depends on the projection angle (the sagittal view was defined to be at position $[0, 0]$). The reason for this being that the SI motion is not visible in near-axial projections. The figure also suggests that there is a certain tolerance to deviate from a perfectly sagittal viewing angle without significantly sacrificing prediction accuracy. This, of course, is an important insight for practical applications where such parameters cannot be chosen freely.

C. Prediction from 1D and 2D Data with Amplitude Gating

In the last scenario we aim at compensating the organ motion for an MR-guided HIFU system under free breathing. As the MR system is used for temperature mapping of the HIFU ablation process, no or only little scan time can be used to track the tumour. In this scenario we thus evaluated the prediction power assuming that the diaphragm can either be tracked with an additional MR-compatible 2D

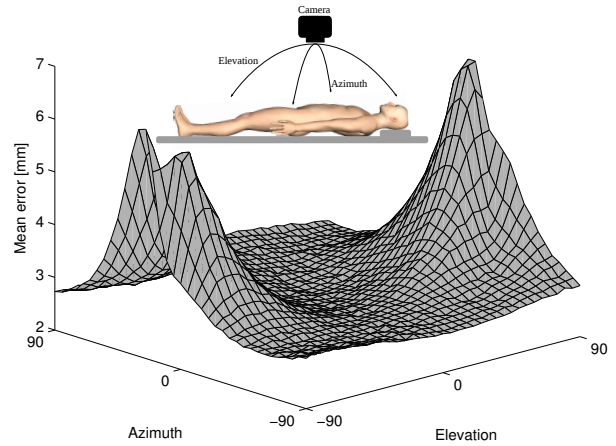


Fig. 5. This surface shows the mean error over all subjects as a function of the camera position. It can be nicely seen that the prediction accuracy drops significantly when *el* approaches the values $+90^\circ$ and -90° .

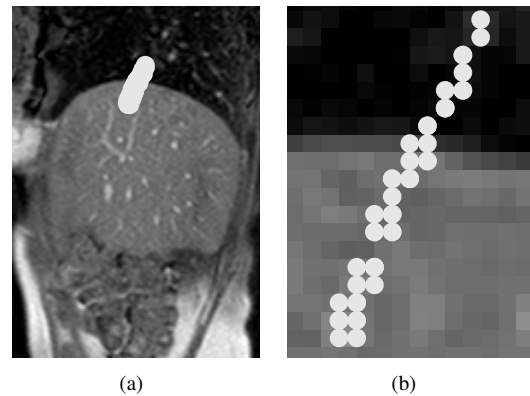


Fig. 6. Selection of tracked organ position using template matching that were used as input for prediction (a), close-up view (b).

ultrasound probe or using a simple and fast pencil beam located on the diaphragm [1]. We applied normalised cross correlation to track one point on the diaphragm over time. Our prediction algorithm is then driven by this offset, which gives a complete 3D position for the liver at every point in time. Figure 6 depicts the navigator slice of one subject together with the tracked positions of the diaphragm. We extracted this data for each subject and obtained results for continuous prediction as shown in Figure 7. The plot shows results for 1D, 2D and 3D predictions. Table I summarise the numbers. In the 1D case we only used the SI component of the signal, which characterises the main mode of motion for the liver. Using two components improves the result slightly. In the 3D case, we used the actual 3D position of the surrogate instead of the tracking results, as in Section III-A. The results show that our model can be used even in cases where only a 1D signal is available, *e.g.* from a respiration belt or an optical tracker, although one has to keep in mind that external signals do not always correlate sufficiently with internal motion [12].

We combined our model-based approach with respiratory gating in order to additionally reduce the prediction error.

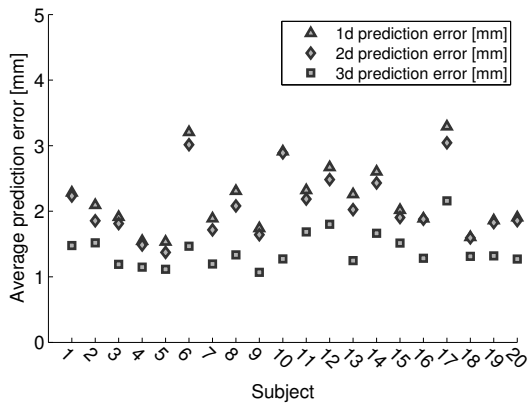


Fig. 7. Average errors for prediction using a single point near the diaphragm in 1D, 2D and 3D.

TABLE I

MEDIAN ERROR AND PERCENTILES FOR LEAVE-ONE-OUT EXPERIMENTS OVER 20 min OF RESPIRATION AND $\eta = 2$.

Input	Median (mm)	Percentiles (mm)		
		25 th	50 th	95 th
1D	2.2	1.7	2.7	3.7
2D	2.1	1.6	2.2	3.5
3D	1.3	0.9	1.4	3.0

Figure 8 depicts results for motion compensation using 1D, 2D and 3D signals (as above) as well as for respiratory gating without compensation. Model-based prediction outperforms gating for any window size. The plot suggests that it is possible to increase the gating window significantly without sacrificing too much accuracy. For example, a gating window of 5 mm results in an average prediction error of 1.8 mm w/o compensation and 1.6 mm in the model-based case. As the gating window is increased to 10 mm, we get an average prediction error of 2.2 mm w/o compensation compared to 1.8 mm (model-based). The error for a 10 mm gating window with model-based motion compensation is thus almost 20% lower than without compensation.

IV. CONCLUSIONS

In this paper we showed how to build statistical motion models and how they can be applied in various realistic scenarios. The prediction power of the motion model was quantitatively evaluated for all scenarios. The results show that organ motion of an unseen subject can be estimated within clinically acceptable margins using a statistical motion model and information from surrogate markers. Additionally, we presented a novel study that compares classical amplitude gating to gating combined with our model-based approach. The results are promising, as they show that our model can improve the accuracy under respiratory gating and can also be used to extend the gating window and thereby reduce treatment times with only a marginal increase in prediction

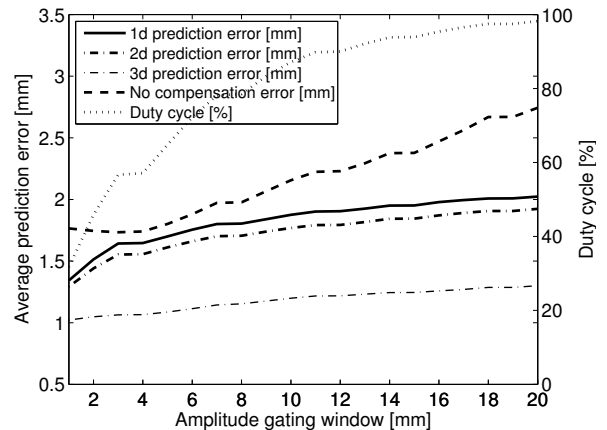


Fig. 8. Comparison of prediction error with amplitude gating. Model-based gating outperforms traditional amplitude gating for any gating window size.

error. In the future work we will use surrogate marker motion data from ultrasound images to predict the tumour location.

ACKNOWLEDGMENT

This work was funded by the Swiss National Science Foundation (SNSF), project CRSII2_127549.

REFERENCES

- [1] P. Arnold, F. Preiswerk, B. Fasel, R. Salomir, K. Scheffler, and P. Cattin. Model-based respiratory motion compensation in MRgHIFU. *IPCAI Proceedings*, 2012.
- [2] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa. The quickhull algorithm for convex hulls. *ACM T. Math. Software*, 22(4):469–483, 1996.
- [3] V. Blanz and T. Vetter. Reconstructing the complete 3D shape of faces from partial information. *Informationstechnik und Technische Informatik*, 44(6):295–302, 2002.
- [4] J. Ehrhardt, R. Werner, A. Schmidt-Richberg, and H. Handels. Statistical modeling of 4d respiratory lung motion using diffeomorphic image registration. *Medical Imaging, IEEE Transactions on*, 30(2):251–265, feb. 2011.
- [5] T. He, Z. Xue, W. Xie, and S. T. C. Wong. Online 4-d ct estimation for patient-specific respiratory motion based on real-time breathing signals. In *Proc. MICCAI 2010*, pages 392–399. Springer, 2010.
- [6] M. J. Murphy. Tracking moving organs in real time. *Semin. Radiat. Oncol.*, 14(1):91–100, 2004.
- [7] F. Preiswerk, P. Arnold, B. Fasel, and P. Cattin. A bayesian framework for estimating respiratory liver motion from sparse measurements. In H. Yoshida, G. Sakas, and M. Linguraru, editors, *Abdominal Imaging. Computational and Clinical Applications*, volume 7029 of *Lecture Notes in Computer Science*, pages 207–214. Springer Berlin / Heidelberg, 2012.
- [8] F. Preiswerk, P. Arnold, B. Fasel, and P. C. Cattin. Robust tumour tracking from 2d imaging using a population-based statistical motion model. In *Mathematical Methods in Biomedical Image Analysis (MMBIA), 2012 IEEE Workshop on*, pages 209–214, jan. 2012.
- [9] 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. Imag.*, 18(8):712–721, 1999.
- [10] H. Shirato, Y. Seppenwoolde, K. Kitamura, R. Onimura, and S. Shimizu. Intrafractional tumor motion: lung and liver. *Semin. Radiat. Oncol.*, 14(1):10–18, 2004.
- [11] M. von Siebenthal, P. Cattin, U. Gamper, A. Lomax, and G. Székely. 4D MR imaging using internal respiratory gating. In *Proc. MICCAI 2005*, volume 3750 of *LNCS*, pages 336–343. Springer, 2005.
- [12] M. von Siebenthal, G. Székely, A. Lomax, and P. Cattin. Systematic errors in respiratory gating due to intrafraction deformations of the liver. *Med. Phys.*, 34(9):3620–3629, 2007.

Chapter 7

Model-Guided Respiratory Organ Motion Prediction of the Liver from 2d Ultrasound

This journal paper, published in *Medical Image Analysis*, describes the system developed together with our project partners of Swiss National Science Foundation (SNSF) Sinergia project CRSII2_127549. We propose methods to spatially and spatio-temporally predict the respiratory motion of the liver from 2d ultrasound images. An MR compatible ultrasound transducer is used to simultaneously acquire ground-truth data using 4D-MRI. Ultrasound is a convenient image modality as it is non-invasive and cheap. However, images often suffer from bad quality due to absorption artefacts from the ribs, particularly in the case of the liver. Sometimes only parts of the liver can therefore be observed. A number of points are tracked in the ultrasound images and our population-based statistical motion model is applied to predict the position of the entire right liver lobe from these sparse measurements. Most of the resulting prediction errors are significantly below 3 mm, which makes the approach interesting for clinical applications.

¹The article is available online at <http://www.sciencedirect.com/science/article/pii/S1361841514000462>.



Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Model-guided respiratory organ motion prediction of the liver from 2D ultrasound



Frank Preiswerk^{a,*}, Valeria De Luca^b, Patrik Arnold^a, Zarko Celicanin^c, Lorena Petrusca^d, Christine Tanner^b, Oliver Bieri^c, Rares Salomir^{d,e}, Philippe C. Cattin^a

^a Medical Image Analysis Center, University of Basel, Switzerland

^b Computer Vision Lab, ETH Zurich, Switzerland

^c Division of Radiological Physics, University of Basel, Switzerland

^d Faculty of Medicine, Radiology, University of Geneva, Switzerland

^e Radiology Department, University Hospitals of Geneva, Switzerland

ARTICLE INFO

Article history:

Received 17 June 2013

Received in revised form 27 March 2014

Accepted 31 March 2014

Available online 13 April 2014

Keywords:

Respiratory motion compensation

Statistical motion model

Spatio-temporal prediction

4D-MRI

Ultrasound

ABSTRACT

With the availability of new and more accurate tumour treatment modalities such as high-intensity focused ultrasound or proton therapy, accurate target location prediction has become a key issue. Various approaches for diverse application scenarios have been proposed over the last decade. Whereas external surrogate markers such as a breathing belt work to some extent, knowledge about the internal motion of the organs inherently provides more accurate results. In this paper, we combine a population-based statistical motion model and information from 2d ultrasound sequences in order to predict the respiratory motion of the right liver lobe. For this, the motion model is fitted to a 3d exhalation breath-hold scan of the liver acquired before prediction. Anatomical landmarks tracked in the ultrasound images together with the model are then used to reconstruct the complete organ position over time. The prediction is both spatial and temporal, can be computed in real-time and is evaluated on ground truth over long time scales (5.5 min). The method is quantitatively validated on eight volunteers where the ultrasound images are synchronously acquired with 4D-MRI, which provides ground-truth motion. With an average spatial prediction accuracy of 2.4 mm, we can predict tumour locations within clinically acceptable margins.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Respiratory organ motion is a complicating factor in the treatment of liver tumours. With the advancement of highly precise tumour treatment options like high-intensity focused ultrasound (HIFU) (Amin et al., 2008; Palussiere et al., 2003) or proton therapy (Greco and Wolden, 2007; Lomax, 1999), non-rigid deformation during breathing introduces a significant amount of uncertainty during treatment using these methods. A simple method to avoid respiratory motion is to completely interrupt breathing while the therapy beam is on (Koshani et al., 2006; Nelson et al., 2005; Pedersen et al., 2004; Balter et al., 2002; Hanley et al., 1999). This method, called *breath-hold*, has the obvious disadvantage that treatment time is prolonged and some patients might not be able to hold their breath at all. Another straight-forward approach is to reduce the extent of breathing by asking the patient to breathe shallowly (Lu et al., 2006; Zhang et al., 2005; Herfarth et al., 2000; Uematsu et al., 2000), which is often not sufficient on its own and

therefore combined with other methods. In *gated* treatment, the beam is only turned on during a certain fraction of the breathing phase, typically near exhalation (George et al., 2006; Jiang, 2006; Keall et al., 2001; Kubo and Hill, 1996; Ohara et al., 1989). Although the aforementioned approaches compensate breathing motion to some extent, they require reproducibility of the organ position for the selected breathing phase (von Siebenthal et al., 2007b) and prolong the treatment time. More importantly, they only compensate for the perpetual breathing motion and as such are completely oblivious to all other modes of organ motion. Thus they are only accurate in a short window of a couple of minutes after patient set-up. Hence, organ motion during radiotherapy continues to be a problem and studies have shown that 4-dimensional treatment planning is important for improved precision in radiotherapy (Shirato et al., 2004).

In order to get past the above limitations, the development of models that compensate organ motion has been an area of intense research during the past decade. Common to many respiratory modelling approaches is the measurement of surrogate data (e.g. displacement of the skin surface, respiratory belt, spirometer or imaging modalities such as ultrasound (US), magnetic resonance

* Corresponding author. Tel.: +41 61 265 96 58.

E-mail address: frank.preiswerk@unibas.ch (F. Preiswerk).

imaging (MRI) or computed tomography (CT)) as the input upon which the model provides an estimate of the motion. Comprehensive overviews on the research field can be found in (McClelland et al., 2013 and Tanner et al., 2012).

A central question for any motion compensation approach is how an accurate respiratory signal is obtained. Some of the approaches rely on a correlation between external signals (e.g. optical tracking, respiratory belt or spirometer) and internal organ motion (He et al., 2010; Simon et al., 2005; Kubo and Hill, 1996). Such approaches suffer from the limitation that external signals do not always correlate well with the actual internal motion of an organ (Hoisak et al., 2006; Murphy, 2004; Vedam et al., 2003). This is because of various reasons, such as respiratory motion not being a rigid process, because of phase-shifts of the different signals and also due to organ drift as elaborated in (von Siebenthal et al., 2007b). A more sophisticated method already used in clinical practise is to implant electromagnetic tracking beacons (EMT) that can be precisely tracked from outside up to sub-millimetre accuracy (Balter et al., 2005). This gives a precise internal respiratory signal, and even more, information about the displacement of the organ of interest itself. Disadvantages of EMT include invasiveness, sensitivity to both metal objects and EM radiation and, furthermore, the high-Z materials used in the beacons have a significant perturbing effect on proton dose distribution (Giebeler et al., 2009). Additionally, in HIFU treatment, the metal can heat up rapidly causing unwanted secondary hotspots. Lastly, the consequences of leaving such markers in the body after the treatment are unstudied.

Most of the motion models described in (McClelland et al., 2013) are custom-built models specific to a patient. In (King et al., 2012), a PCA-based motion model is proposed to reconstruct 3d lung MRI over time by registering a 3d reference image at exhalation to a 2d navigator image through optimisation of the model parameters. This is similar to the approach presented here, but their model is subject-specific and the approach relies on the availability of intra-operative 2d MRI. While tailor-made models can be highly precise, the acquisition of patient-specific data for model construction is time-intensive, costly and therefore not always feasible in clinical practise. As mentioned in (McClelland et al., 2013), cross-population approaches are a promising research field that can potentially eliminate the need to acquire subject-specific images. A number of such cross-population approaches have been proposed over the past few years (Preiswerk et al., 2012b; Ehrhardt et al., 2011; Arnold et al., 2011; He et al., 2010; Ehrhardt et al., 2010; Fayad et al., 2009; Ehrhardt et al., 2008; Klinder et al., 2009; Sundaram et al., 2004). In (He et al., 2010), 4D motion of lungs is modelled using Kernel PCA. A support vector machine is trained to model the relation between motion of fiducial markers on the lower abdomen/chest and the coefficients of the K-PCA. However, a correlation between organ motion and the motion of an external surrogate may become unreliable over time in presence of organ drift (von Siebenthal et al., 2007b). In (Ehrhardt et al., 2008) it is shown that a mean motion model built from 4D-CT images of the lung can be used for predicting lung and tumour motion of a novel subject without knowledge of the individual breathing dynamics. In (Klinder et al., 2010), a statistical motion model is used to predict lung motion by learning multivariate linear regression between the parts to be predicted and a 4D-CT training set of 10 patients. In (Ehrhardt et al., 2011), diffeomorphic nonlinear intensity-based registration and the Log-Euclidean framework is used to build a motion model from thoracic 4D-CT lung data sets of 17 patients. The respiratory cycle is discretised into four states and 4D-CT images of the patient are required for estimating a scaling factor for breathing depth. A prediction accuracy of 3.3 mm on average is reported. In contrast to these data-driven approaches, there is also work on physically-based

simulations of respiratory organ motion using mass-spring or finite-element methods (Eom et al., 2010; Werner et al., 2009). While such approaches are promising, knowledge of the underlying physical mechanism and tissue properties is required, which is still a limiting factor.

In this study, we demonstrate the applicability of a population-based statistical respiratory motion model of the liver that is driven by 2d US imaging. A Bayesian reconstruction algorithm computes the most likely prediction of the liver's location based on fiducials tracked in the US images (De Luca et al., 2012). This way, the lesion to be tracked itself need not be visible in the US images. In addition, we show how spatial prediction can be extended with a temporal predictor to estimate the position of the liver in the future using an artificial neural network (ANN). This is essential for practical applications to compensate for the time required to process the US images and the lag for refocusing the treatment beam. Simultaneously to the US images, we acquired four-dimensional magnetic resonance images (4D-MRI) of the liver (Petrusca et al., 2013) for validation. This allows for a quantitative evaluation of the prediction accuracy over long time scales.

Our statistical motion model is built beforehand from a representative set of healthy subjects using displacement fields obtained from 4D-MRI images (Preiswerk et al., 2012a). These subjects are distinct to the subjects used for prediction. For the latter eight subjects, we acquired the same kind of 4D-MRI images (simultaneously to the 2d US data) for validation; however, we do not assume any 4D-MRI to be available during prediction.

The contributions of this paper are twofold. First, we combine robust feature tracking in 2d US images (De Luca et al., 2012) for driving a population-based statistical motion model (Preiswerk et al., 2012a) for spatial and spatio-temporal prediction of respiratory liver motion. Second, we combine 4D-MRI and US imaging (Petrusca et al., 2013) for validating our results. Although some parts have previously been published, this is the first study that describes the integration of these components. In particular, (De Luca et al., 2012) is not using a statistical model, (Preiswerk et al., 2012a,b) is not using US images but only a virtual breathing signal extracted from the 2d navigator slice of the 4D-MRI sequence. The novel developments reported here are part of a long-term project and represent the final integration of methods and subsystems.

2. Materials

This section describes the hardware setup used to acquire simultaneous 4D-MRI/US and briefly explains how a population-based statistical motion model is built (from a separate set of 4D-MRI data). A schematic overview of the system is depicted in Fig. 2.

2.1. MR-compatible ultrasound

Ultrasound images were acquired using an Acuson clinical scanner (Antares, Siemens Medical Solutions, Mountain View, CA, USA) equipped with a 256 phased array transducer. The device was modified by the manufacturer to avoid magnetic materials, thus ensuring passive MRI compatibility (Petrusca et al., 2013). A customised holder for the US probe (Fig. 1 2(a)) was fabricated using stereolithography (electrically insulating resin) in order to allow an optimal window for the US beam. The probe was rigidly attached inside this holder. To prevent noise at the radiofrequency detection range of the MRI scanner, electromagnetic shielding was required. Therefore, the holder and its transmission line (7 m) were entirely shielded using copper and aluminium coating, respectively. The entry gate of the beam-former was adjusted to

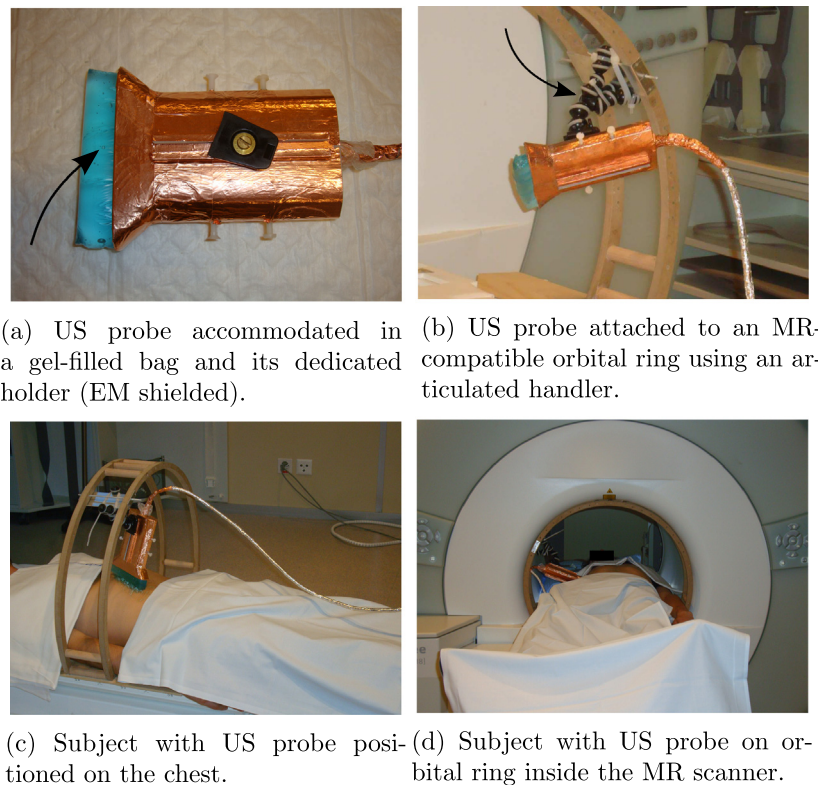


Fig. 1. Overview of the experimental setup.

compensate for the increased capacitance of the longer transmission line. Acoustic coupling between the patient skin and the US probe was assured by a plastic bag filled with standard ultrasonic gel. The bag was attached to the active part of the US transducer. Besides acoustic coupling, the gel-bag ensured respiratory motion decoupling. Throughout respiration, the bag deforms while the US probe remains stationary. Such motion decoupling is fundamental, since our experiments rely on a fixed co-registration between the US probe and the MRI scanner. Furthermore, the gap between the skin and the probe ensured that the device remains sufficiently distant from the subject's body to avoid any residual susceptibility artefacts. The US probe located inside the shielded holder was attached to an MR-compatible orbital ring using an MR-compatible articulated handler (Fig. 1(b)). To facilitate the positioning of the US probe on the ring, five degrees of freedom of the holder were available with this setup: transversal translation of the orbital ring on the MRI bed, circular revolution of the handler following the orbital ring's perimeter, two rotations in the articulated handler and one final translation of the shielded holder parallel to the base of the articulated handler, to control the distance from the US imaging probe to the patient's skin. Figs. 1(c) and (d) show a subject on the MR bed with the US probe positioned on the chest, ready for data acquisition. More details on the US shielding for MRI and a detailed analysis on RF interference and the US imaging-induced changes of the SNR in MRI are reported in (Petrusca et al., 2013).

Ultrasound sequences of 14–17 Hz temporal and 0.28–0.4 mm spatial resolution were acquired to observe the liver motion for eight subjects. For six subjects, the transducer was placed on the intercostal space between ribs number 6 and 8 on the right side and the image plane was aligned to the longitudinal axis as much as possible to minimise out-of-plane motion during respiration. For the other two subjects, the probe was placed frontally on the abdomen, acquiring a near-sagittal image. Completely avoiding

Table 1

Overview of the datasets acquired for motion prediction in this study.

Subject	1	2	3	4	5	6	7	8
Gender	m	m	m	m	f	f	f	m
Age	46	30	37	33	28	30	30	32
No. of phases	840	840	840	840	840	840	780	784
No. of cycles	97	61	71	94	96	52	58	63
<i>Image resolutions (mm)</i>								
4D-MRI inplane	1.82	1.82	1.82	1.82	1.82	1.82	1.82	1.82
4D-MRI slice distance	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
2d Ultrasound	0.40	0.40	0.37	0.36	0.36	0.28	0.42	0.40

any transversal component in the image is not possible because the acoustic window is restricted by the intercostal space of the ribs. Table 1 summarises the properties of the acquired simultaneous 4D-MRI/US data.

2.2. 4D-MRI and 3D image registration

Using the previously described MR-compatible US setup, we acquired simultaneous 4D-MRI/US images (von Siebenthal et al., 2007a; Petrusca et al., 2013) of eight subjects to obtain a respiratory signal for prediction (from US) and 4d ground-truth for validation (from 4D-MRI). The data was recorded over 4–5 blocks of 5.5 min each, with a break of 1 min between blocks. Image acquisitions were performed on a 1.5 T whole-body clinical scanner (MAGNETOM Aera, Siemens Medical Solutions, Erlangen, Germany). A combination of phased-array coils (10 active elements) was used, aiming to provide a homogeneous signal from the whole abdomen. High vessel-tissue contrast, necessary for the detection of the organ position, was generated with a balanced steady-state free precession (bSSFP) sequence giving bright vessel signal due to

its inherent flow compensation, its T2/T1-weighted contrast, and inflow-related signal enhancement. The main acquisition parameters were: flip angle 78°, TE/TR 1.38/3 ms, in-plane resolution 1.82 mm isotropic, 26–30 slice positions, slice thickness 6 mm with slice gap –2 mm, image matrix 192 × 192, SENSE factor 2. Given the above acquisition parameters, a temporal resolution of 5.2 Hz is achieved. Because of the interleaved acquisition scheme in 4D-MRI where navigator/data pairs are subsequently acquired, this results in a temporal resolution of 5.2 Hz/2 = 2.6 Hz for the navigators and the data slices, respectively.

4D-MRI works through interleaved image acquisitions of a fixed-position navigator slice and a moving data slice that samples the organ of interest equidistantly, in our case sagittally. From the embracing navigators, the breathing state is determined to generate 4d-reconstructions from the data slices. This is done using the NiftyReg open-source registration software (Modat et al., 2010), which is a fast implementation of the B-spline-based free-form deformation algorithm presented in (Rueckert et al., 1999). First, a rigid transformation is optimised that serves as the initialisation for the non-rigid registration and optimisation is performed in a multi-scale fashion. This gives a 2d displacement field for each navigator image. The 4d reconstruction is based on the assumption that corresponding phases at different times are subject to similar overall deformation of the liver. Since every data slice is encompassed by a preceding and a subsequent navigator image, data slices belonging to the same breathing state can be retrieved by comparing the deformation fields of their encompassing navigator frames. When enough data has been acquired, a 3d image can be retrospectively reconstructed for each encapsulating navigator image pair, thereby creating a 4d image sequence. In our experiments, there were no significant artefacts in the 4D-MRI volumes due to the fact that the individual 2d slices that make up a 3d stack were acquired at different points in time.

From the reconstructed 4D-MRI image sequence, the first exhalation image is chosen as the exhalation master and is manually segmented. This image is subsequently registered to all images of the 4D sequence using the same registration scheme as described previously for the 2d navigators. A common problem for registration of moving organs are discontinuities occurring at the border between the abdominal wall and the organ sliding along it in superior–inferior direction. Although interesting discontinuity-preserving registration algorithms have recently emerged (Kiriyanthan et al., 2012), they are computationally too expensive. We circumvent this problem by masking the liver in the exhalation master image and only evaluate the similarity measure (normalised cross correlation) within this masked region. All registration tasks were performed on a cluster consisting of 70 CPUs at ~2.5 GHz and between 4 GB and 8 GB memory each.

2.3. Statistical model of respiratory motion

To model and predict the respiratory liver motion in our experiments, we use the population-based statistical motion model from (Preiswerk et al., 2012b). This model consists of liver motion data from 20 healthy individuals (9 females, 11 males, average age 30, range 17–75, similar protocol as described in Section 2.2 with significant difference only in flip angle 70°, SENSE factor 1.7, temporal resolution of 5.8 Hz/2 = 2.9 Hz). The subjects contained in this model are disjoint to the eight simultaneous 4D-MRI/US datasets described previously. In this section, we briefly elaborate on how the model is built.

2.3.1. Establishment of inter-subject correspondence

To build a population-based model, correspondence must be established among all datasets. For each subject, four landmarks on the liver surface are manually selected on each slice of the master exhalation volume. The user-selected landmarks are then projected to their closest point on the contour of the segmented liver, dividing the contour into four segments. Each segment is equidistantly subsampled to obtain a final point set for each slice (Fig. 3(a)). Additionally, the points are also equidistantly subsampled in the direction of slice acquisition. This is done because the liver covers a different number of slices per subject. The result of these remeshing steps is a master shape for each subject consisting of 900 surface points that are in correspondence between all subjects. In a next step, we establish correspondence for points *inside* the shape among all subjects. One approach would be to also (manually) select corresponding anatomical structures within the liver. However, since the vessel branch structure can be very different even for homozygotic twins, we prefer an automatic approach that interpolates the correspondence from the surface of each subject to the interior. To do this, we calculate the mean of all master shapes through a partial (no scaling) generalised Procrustes analysis (GPA) (Dryden and Mardia, 1998). The mean is then embedded within an isotropic grid of 10 mm spatial resolution, see left column in Fig. 3(b). Finally, this grid is warped to all 20 master shapes using piece-wise affine transformations based on a tetrahedrisation of the mean grid, as described in (von Siebenthal et al., 2007a; Barber et al., 1996). For each subject, this results in a set of $m = 2142$ corresponding surface and interior model points. The right column of Fig. 3(b) depicts this for one subject.

2.3.2. Principal Component Analysis (PCA)

The same non-rigid 3d registration as previously described is also performed on the 20 subjects for model construction. The displacement fields are extracted at corresponding grid locations. For

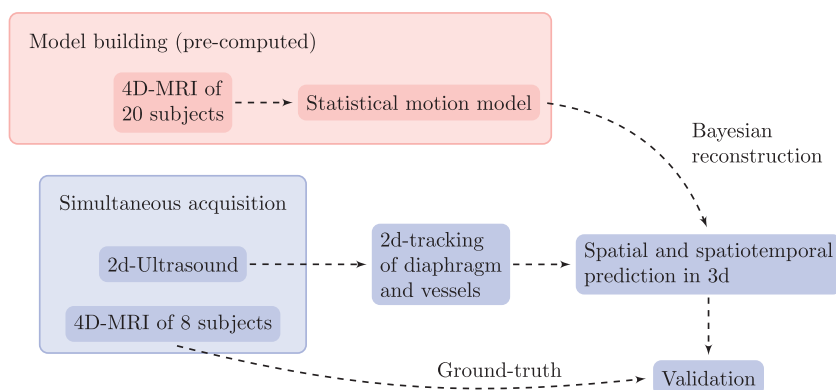


Fig. 2. Overview of the involved data acquisitions and processing. The statistical motion model is pre-computed and does not contain the subjects used during simultaneous acquisition of 4D-MRI and 2d US.

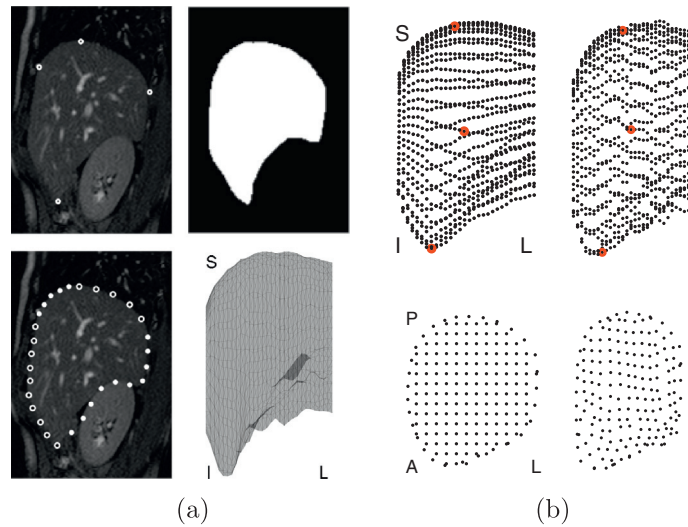


Fig. 3. (a) Manually selected landmarks for correspondence on a selected sagittal slice (top left), Manually defined mask (top right), resampled contour based on mask and landmarks (bottom left), resulting 3d mesh (bottom right). (b) Coronal view of the mean exhalation shape with isotropic grid and three points highlighted (top left), master exhalation shape of a selected subject with the same corresponding points highlighted (top right). Transversal slices of the shapes showing internal grid points for the above examples (bottom row).

every subject s and time step t , the liver is represented by a vector $\mathbf{v}_s(t) = (x_1(t), y_1(t), z_1(t), \dots, x_m(t), y_m(t), z_m(t))'_s \in \mathbb{R}^{3m}$.

We discard the shape information and only consider the relative displacement $\mathbf{x}_s(t)$ to the exhalation master $\check{\mathbf{v}}_s$,

$$\mathbf{x}_s(t) = \mathbf{v}_s(t) - \check{\mathbf{v}}_s. \quad (1)$$

For each subject, we select the respiratory states of five breathing cycles sampled uniformly across the complete duration. We also select the five deepest cycles. Let $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ be the matrix of all selected samples over all subjects and $\boldsymbol{\mu} = \frac{1}{n} \sum \mathbf{x}_i$ the sample mean. Singular value decomposition (SVD) of the covariance matrix $\Sigma = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T$ yields

$$\Sigma = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T, \quad (2)$$

the orthonormal matrix of principal components $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_{n-1})$ and their corresponding eigenvalues $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_{n-1})$ that represent the standard deviation σ_i of the principal components in descending order ($\sqrt{\lambda_i} = \sigma_i$). We can transform (and thus decorrelate) the data by subtracting the mean, followed by a projection into model space,

$$\mathbf{c}_s(t) = \text{diag}(\sigma_i^{-1})\mathbf{U}^T(\mathbf{x}_s(t) - \boldsymbol{\mu}). \quad (3)$$

Eq. (1) can be written in terms of the model coefficient vector \mathbf{c}_s ,

$$\mathbf{x}_s(t) = \text{diag}(\sigma_i)\mathbf{U}\mathbf{c}_s(t) + \boldsymbol{\mu}. \quad (4)$$

The model can be efficiently compressed by omitting eigenvectors in \mathbf{U} that have low corresponding eigenvalues. A typical threshold is to retain 95% of their cumulative sum. Applying the model for motion prediction boils down to finding a suitable coefficient vector $\mathbf{c}(t)$ that optimally represents the observed motion of a set of points extracted by US tracking. In Section 3.3 we show how this is achieved.

3. Methods

In this section, we describe the algorithm used to track points in US images (Section 3.1), the co-registration between MR and US (Section 3.2) and the Bayesian algorithm used for spatial prediction (Section 3.3). Lastly, we also present a method to perform spatio-temporal prediction in Section 3.4.

3.1. Ultrasound tracking

The required temporal resolution of the US tracking to keep the system lag low disqualifies traditional registration algorithms. Our aim is to extract the motion of multiple fiducials in the images, rather than computing the displacement field of the whole organ under investigation. Therefore, we use the tracking method described in (De Luca et al., 2012). This method exploits the redundancy within the images in the case of quasi-repetitive motion scenarios, such as breathing, and consists of a training and an application phase. During training, images of a previously acquired sequence are registered and the relationship between the image appearance and the spatial transformation is learned. For each image in the application phase, the most similar images in the training set are selected for estimating the associated displacements of points of interest. Registration and update of the training data is only performed for outliers.

3.1.1. Training phase

During the training phase, an initial US sequence of 1 m length is acquired, starting at exhalation. The US images $I(t_i)$ are registered to the initial image $I(t_0)$ by optimising the parameters of a local affine transformation (six degrees of freedom) with respect to normalised cross correlation. The transformation is defined for a region enclosing the tracked structure, e.g. a region on the diaphragm or a vessel. The points to be tracked \mathbf{p} are manually selected in $I(t_0)$. The registration result from the previous image pair ($I(t_{i-1})$ to $I(t_0)$) is used as an initial estimation. The image appearance is stored by embedding $I(t_i)$ into a PCA space $\mathbf{S}(t_i)$. \mathbf{S} and the corresponding registration results (i.e. \mathbf{p}) are stored $\forall t_i$ in the training set.

3.1.2. Application phase

During a real-time procedure, new images are continuously acquired. Given the current image $I(t)$, we want to find the most similar image in the training set for obtaining a prediction of the current displacements. After projecting $I(t)$ into the PCA space \mathbf{S} , we determine the K -nearest neighbours (K -NN) to $\mathbf{S}(t)$ in the training set. In detail, we find t_{m_k} with the smallest distances

$$d_k = \|\mathbf{S}(t) - \mathbf{S}(t_{m_k})\|, \quad (5)$$

with $1 \leq k \leq K$ and K the number of nearest neighbours ($K \in [2, \dots, 4]$). The distances are sorted so that $d_k \leq d_{k+1}$. Having determined the K -NN from the training set, we retrieve the K associated positions $\hat{\mathbf{p}}(t_{m_k})$ per tracked point. The prediction of $\mathbf{p}(t)$ is then obtained by linearly interpolating $\hat{\mathbf{p}}(t_{m_k})$,

$$\mathbf{p}(t) = \sum_{k=1}^K \left(1 - \frac{d_k}{\sum_k d_k}\right) \hat{\mathbf{p}}(t_{m_k}). \quad (6)$$

As completely repetitive motion is very unlikely, a mechanism to cope with previously unseen variations (outliers) and adapt to these (e.g. drift of exhalation position) is incorporated. In order to detect such outliers, we establish a threshold on the image similarity based on the statistics of the training data. Specifically, for each training image we find its nearest neighbour to the remaining training samples and calculate the minimum Euclidean distance d_1 according to Eq. (5). A threshold Θ is then set to the 95th percentile of these distances. During the application phase, if $d_k(t) > \Theta$ then $I(t)$ is categorised as outlier and registration is computed as described above. Finally, we update the training database by adding the registration result and the PCA description of the outlier, and by recalculating Θ . Fig. 4 depicts a US image with typical points of interest as well as their tracking results over five cycles. The available choice of trackable structures is significantly different among subjects. It is preferable to have at least a good view on the diaphragm for tracking, as the breathing motion induced by it dominantly defines the respiratory motion of the liver.

3.2. Co-registration of modalities

The application of our prediction algorithms requires precise co-registration of the involved image modalities. As the organ motion is observed in 2d US images, the tracked points must be transformed from the 2d US image plane to 3d positions in the MRI coordinate frame. To establish this correspondence, a breath-hold image acquisition of the abdomen is performed (3d T1-weighted fat-saturated volumetric interpolated breath-hold examination (VIBE), TE/TR 1.96/4.36 ms, flip angle 10° , image resolution $1.25 \text{ mm} \times 1.25 \text{ mm} \times 2 \text{ mm}$). During the acquisition of this sequence, all volunteers are instructed to suspend respiration at end expiration. In this image, the MR-compatible case of the US transducer is visible, as highlighted in Fig. 5(a). The orientation of the US plane is extracted through manual inspection of the apnea image volume (Fig. 5) and results in a 3d rotation $R_{3d} \in \mathbb{R}^{3 \times 3}$. Furthermore, a rigid 2d transformation between the US image and its corresponding MRI slice is obtained through optimisation from a number of manually placed landmarks on both images and results

in rotation $R_{2d} \in \mathbb{R}^{2 \times 2}$ and translation $T_{2d} \in \mathbb{R}^{2 \times 2}$. Finally, the US points $\mathbf{p}(t)$ can be mapped to the common 3d coordinate space,

$$\mathbf{p}_{3d}(t) = R_{3d} T_{13d} T_{2d} R_{2d} \cdot \mathbf{p}(t), \quad (7)$$

where T_{13d} maps $\mathbf{p}(t) = (p_x, p_y)'$ to its position $(p_x, p_y, p_z)'$ on the (manually identified) US slice in the rotated apnea volume (Fig. 5(b)). Fig. 5(d) depicts the final alignment of US to MRI of a selected subject.

In a next step, correspondence to the model is established using the same method as described in Section 2.3.1 (manual segmentation on each MR slice of the exhalation master, selection of landmarks on each slice). Via Eq. (7), an exhalation position $\hat{\mathbf{p}}$ of each tracked point in US is then transformed to $\hat{\mathbf{p}}_{3d}$ and assigned to its closest index k in the model. This mapping is later used to assign the observed motion in US to model vertices during prediction.

The position of the US plane can only be aligned with the direction of respiratory organ motion to a certain extent. A good position of the transducer is the posterior abdominal region, with the US plane oriented sagittally. However, in cases where this leads to significant pressure on the transducer due to respiration, which causes the US transducer to move, we positioned the US transducer in the right lateral region of the chest. The limitation of this is the smaller acoustic window through the intercostal region between the ribs. Together with the non-rigid nature of respiratory motion, the consequence is that out-of-plane motion in the 2d US sequence is unavoidable. To correct for this, every tracked point $\mathbf{p}_{3d}(t)$ (lying on a plane) is projected onto the mean motion direction of its assigned motion model vertex. For a motion vector $\Delta \mathbf{p}_{3d}(t) := \mathbf{p}_{3d}(t) - \hat{\mathbf{p}}_{3d}$ and the mean motion $\boldsymbol{\mu}_k$ of its associated vertex k in the motion model, the projection is given by

$$\Delta \hat{\mathbf{p}}_{3d}(t) = \cos(\alpha) \cdot \|\Delta \mathbf{p}_{3d}(t)\| \cdot \frac{\boldsymbol{\mu}_k}{\|\boldsymbol{\mu}_k\|}, \quad (8)$$

where α is the angle between $\Delta \mathbf{p}_{3d}(t)$ and $\boldsymbol{\mu}_k$. Clearly, this is only an approximation of the true out-of-plane motion but it is the most probable out-of-plane motion given the model. This projection led to better experimental results compared to feeding the prediction with unfiltered motion directly from the US plane. Furthermore, the simulations performed in (Preiswerk et al., 2012b) have revealed that circular revolution around the longitudinal axis of the patient has only little effect on prediction accuracy using a statistical motion model.

3.3. Spatial prediction

For each time step t , the goal is to estimate a displacement field $\mathbf{x}(t) \in \mathbb{R}^{3m}$ based on the tracked US points. For this task, we apply

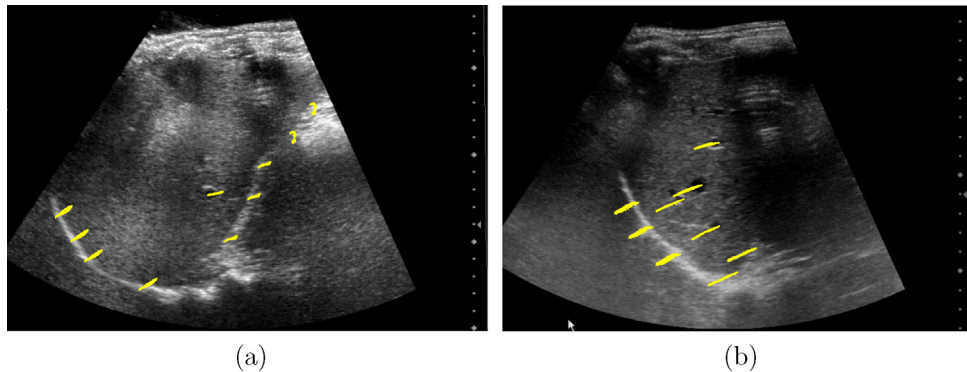


Fig. 4. Ultrasound tracking of five cycles for two subjects. (a) tracked points of subject 1 are distributed mainly along the liver surface, as the available acoustic window did not expose many vessel structures inside the organ. (b) trajectories for subject 4 could be extracted at the diaphragm and also for vessels within the liver. The acoustic shadow occluding about half of the liver is a typical artefact due to absorption by the ribs.

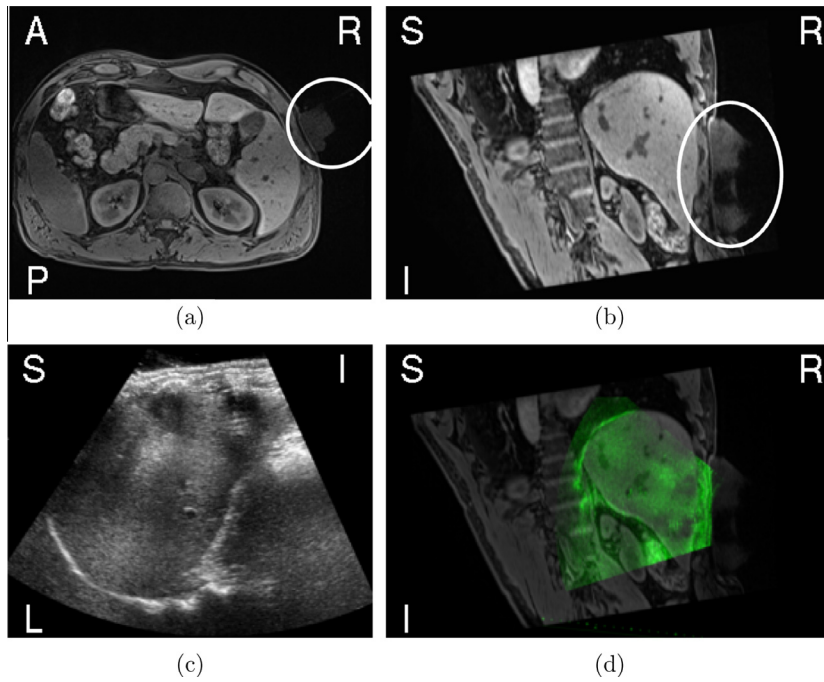


Fig. 5. (a) Apnea slice with transducer highlighted on the right. (b) Slice through the apnea image corresponding to the US plane with transducer highlighted. (c) Ultrasound image. (d) MR and US images aligned and overlaid.

the Bayesian approach described in (Blanz and Vetter, 2002), which computes the most probable model reconstruction given partial data through a maximum a posteriori (MAP) estimation of the model coefficients. The algorithm provides an optimal reconstruction even in the presence of noisy measurements. Let $\mathbf{r}(t) := \Delta \mathbf{p}_{3d}(t)$ be a sparse observation for which we want to estimate the complete displacement field $\mathbf{x}(t)$ of the organ and let \mathbf{L} be the process that generates our sparse observations. From Section 2.3.2 and Eq. (3) we know this is equivalent to finding the model coefficients $\mathbf{c}(t) = \text{diag}(\sigma_i^{-1}) \mathbf{U}^T (\mathbf{x}(t) - \boldsymbol{\mu})$ that best describe the observation. Using Bayes' rule, this can be formulated as a MAP problem,

$$p(\mathbf{c}|\mathbf{r}) = v \cdot p(\mathbf{r}|\mathbf{c}) \cdot p(\mathbf{c}), \quad (9)$$

(where we have omitted parameter t and index s without loss of generality), with normalisation constant v . It can be shown (Blanz and Vetter, 2002) that this quantity is maximised if the cost function

$$E(\mathbf{c}) = \|\mathbf{Q}\mathbf{c} - \mathbf{r}\|^2 + \eta \cdot \|\mathbf{c}\|^2 \quad (10)$$

is minimal. $\mathbf{Q} = \text{diag}(\sigma_i) \cdot \mathbf{L}\mathbf{U}$ is the reduced and scaled model matrix of the PCA (from Eq. (2)) and η is a regularisation factor which can be adjusted to the amount of noise in the observation. The solution to this problem is then given by

$$\mathbf{c} = \widehat{\mathbf{V}} \text{diag} \left(\frac{w_i}{w_i^2 + \eta} \right) \widehat{\mathbf{U}}^T \mathbf{r} \quad (11)$$

where $\mathbf{Q} = \widehat{\mathbf{U}}\mathbf{W}\widehat{\mathbf{V}}^T$ is obtained by SVD and w_i the diagonal entries of \mathbf{W} . A more detailed elaboration of this result is available in (Blanz and Vetter, 2002). Note that all terms of Eq. (11) apart from the observation \mathbf{r} itself, can be pre-computed, *i.e.* no optimisation is necessary during runtime.

3.4. Temporal prediction

Spatial prediction alone is not sufficient for real-world tumour tracking. A tracking algorithm typically also has to *temporally* predict into the future. A robotic radiosurgery device must continu-

ously observe the patient motion, compute the new robot position and mechanically adjust its position. Furthermore, our spatial motion prediction algorithm described above also adds to the system latency because of the involved computations at each time step. All these components together make up the co-called system latency. For example, the system latency of the CyberKnife system (Accuray, Inc., Sunnyvale, CA/USA) is reported to be $t_\Delta = 115$ ms (Seppenwoolde et al., 2007). Thus, at time t , the displacement field $\mathbf{x}(t + t_\Delta)$ must be estimated. Temporal prediction can be integrated in a straight-forward way with our spatial prediction approach from Section 3.3. In principle, any temporal prediction algorithm can be used to estimate the location of the US points at $t + t_\Delta$. Then, the spatial reconstruction is computed from the estimated points $\widehat{\mathbf{p}}(t + t_\Delta)$ instead of $\widehat{\mathbf{p}}(t)$. In (Krauss et al., 2011) it was shown that artificial neural networks (ANN) are well suited for temporal prediction and slightly outperform other established algorithms. Therefore we rely on ANNs for temporal prediction in this study.

For each subject, a function $f: X \rightarrow Y$ that maps a set of n input observations to a desired output value is learned using an ANN. The network is trained using current and past US tracker positions over a signal history length of $l = 3$. Specifically, the average histories $\bar{x}(t) = \frac{1}{n-1} \sum_{i=1}^n \sum_{j=0}^l (x_i(t-j) - x_i(t-j-1))$ (and similar for $\bar{y}(t)$) are presented to the network as input. The output values are the set of desired future positions, $\{x_1(t + t_\Delta), y_1(t + t_\Delta), \dots, (x_n(t + t_\Delta), y_n(t + t_\Delta))\}$, and the network is trained using the Levenberg-Marquadt backpropagation algorithm (Hagan and Menhaj, 1994) on a sequence of 20 s. The network is depicted in Fig. 6.

The choice of an ANN here shall only be exemplary. Other possible temporal predictors that can be integrated, such as the Kalman filter, linear prediction and many more are described in (Sharp et al., 2004; Welch and Bishop, 1995; Putra et al., 2008; Ernst et al., 2013).

4. Experimental results

We have evaluated our approach on eight healthy subjects (no known liver pathologies), consisting of three female and five male

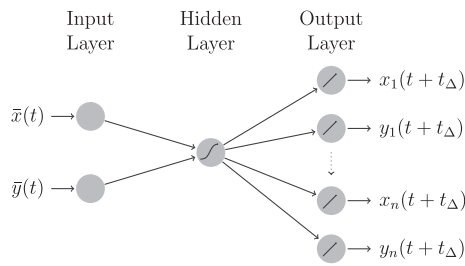


Fig. 6. Overview of the artificial neural network (ANN) trained for temporal prediction. A sigmoidal transfer function is used in the hidden layer and linear transfer functions are applied on the output layer.

individuals between 28 and 45 years of age (see Table 1). Each subject was carefully placed on the MRI bed, paying special attention to a precise alignment of the longitudinal axis of the body and the scanner. The time needed for patient-setup was 60 m for the first trials, mainly due to the time necessary to find and fix a suitable position for the US transducer. However, this quickly reduced to roughly 30 m after gaining experience in the process. Between four and five blocks of 4D-MRI data, with a duration of 5.5 m each, were acquired. Simultaneous 4D-MRI and US imaging was performed during the first block. Using data from all 4–5 acquisition blocks, 4D-MRI reconstructions for the first block were computed. This data was then registered as described in Section 2.2 to obtain automatically-generated ground-truth for the experiments. The US images were registered offline using the approach described in Section 3.1 to produce the input for model-based prediction as described in Section 3.3 and 3.4. As mentioned in Section 2.3, the model was built from a different set of 4D-MRI acquisitions of 20 healthy subjects. For spatio-temporal prediction, an ANN was connected ahead of the model for prediction as described in

Section 3.4. Multiple ANNs (one for each t_Δ) were trained for each of the eight subjects based on tracking results from a short US sequence of 1 m duration, acquired prior to the sequence used for prediction. All predictions described in this section were computed for the states of the 4D-MRI sequence of the first block of 5.5 m, i.e. $t = (0, \dots, n - 1) \cdot 1/2.6$ Hz, where n is the number of 4D-MRI frames (between 784 and 840 for all eight subjects).

4.1. Evaluation of ultrasound tracking results

First, the results from the US tracking algorithm were quantitatively evaluated by computing the statistics of the tracking error with respect to manual annotation of the vessel centres, which an observer manually annotated. The annotation was performed on 10% of the images in the application phase (approximately 450 frames). As a result, the mean tracking error is 0.87 mm, standard deviation 0.47 mm. Additionally, the algorithm is real-time, with a mean computational time of 9 ms per frame.

4.2. Evaluation of spatial prediction

Spatial motion prediction was performed for all subjects according to Section 3.3. Table 2 summarises the results. The regularisation factor was determined empirically and set to $\eta = 1.0$. Let $x(t, i)$ be the position of model vertex i at time t . Our error measure is the Euclidean distance between the prediction and its ground-truth,

$$e(t, i) = \|\tilde{x}(t, i) - x(t, i)\|. \tag{12}$$

Table 2 summarises the results. The mean of $e(t, i)$ is in the range of 1.8 mm and 3.1 mm for all eight subjects. The last row of Table 2 additionally gives the average Dice coefficient between $\tilde{x}(t)$ and $x(t)$, which is consistently higher or equal to 0.9.

Table 2

Results for spatial prediction ($t_\Delta = 0$) over approx. 5 min. Columns 4–6 give percentiles of $e(t, i)$. ‘Amplitude’ gives the breathing amplitude averaged over all model vertices (measured on the ground-truth) and ‘Dice’ gives the mean Dice coefficient over t between the reconstructed liver and its ground-truth.

Subject	Error $e(t, i)$ (mm)						Amplitude (mm)		Dice μ
	μ	σ	p[50]	p[95]	p[99]	max	μ	max	
1	2.0	1.7	1.7	4.3	8.7	31.0	6.7	8.1	.93
2	1.8	1.2	1.6	4.0	5.4	16.5	10.8	13.1	.94
3	2.5	2.1	1.7	6.9	9.6	27.7	10.1	15.1	.92
4	3.1	2.2	2.5	7.3	10.0	32.8	11.8	15.8	.93
5	2.4	1.2	2.3	4.5	6.0	19.0	5.3	6.5	.90
6	2.6	1.7	2.3	5.8	7.5	20.3	18.5	20.5	.92
7	2.3	1.9	1.6	6.1	8.3	26.1	16.7	23.9	.93
8	2.3	1.4	2.0	4.9	6.5	23.7	11.2	14.7	.92
Mean	2.4	1.7	2.0	5.5	7.7	24.6	11.4	14.7	.92

Table 3

Experiments using manually-annotated ground-truth: target registration error (TRE) for validation of the registration results (columns 2–7), spatial prediction error ($t_\Delta = 0$) (columns 8–13).

Subject	TRE (mm)						Error $e(t, i)$ (mm)					
	Observer 1			Observer 2			Observer 1			Observer 2		
	μ	σ	max	μ	σ	max	μ	σ	max	μ	σ	max
1	1.3	0.8	4.7	1.1	0.9	5.1	2.2	1.7	7.4	1.8	1.4	6.4
2	1.3	0.9	6.2	1.5	1.1	6.2	1.5	0.8	4.0	1.6	0.9	4.2
3	1.6	0.4	5.0	1.5	0.5	4.7	2.3	1.0	5.7	2.0	1.1	4.9
4	1.3	0.8	7.3	1.3	0.8	6.8	3.8	2.9	11.6	3.2	2.6	10.5
5	0.9	0.4	3.1	0.8	0.5	3.7	1.9	0.9	4.0	2.2	1.1	5.5
6	2.1	1.2	10.4	2.0	1.1	8.3	2.0	1.2	5.8	2.3	1.5	6.1
7	2.2	1.7	11.9	2.4	1.7	11.8	2.1	1.3	7.5	2.0	1.4	8.2
8	1.8	1.0	6.2	1.5	0.8	5.9	1.5	0.9	3.8	2.2	1.3	5.8
Mean	1.6	0.9	6.8	2.1	1.3	6.2	1.5	0.9	6.6	2.2	1.4	6.5

Although the automatically-generated ground-truth enables to evaluate results over an unprecedented amount of data, we additionally validated the registration as well as the spatial prediction results on manually-labelled data. Therefore, we have asked two medical experts to label two landmarks on each MR navigator slice over five respiratory cycles for each subject. More specifically, they were asked to label a clearly visible vessel structure near the centre of the organ (in S/I and A/P directions) and one position near the inferior end of the liver. This procedure was performed in 2d because the out-of-plane motion as well as the corresponding image-resolution (sagittally) is very small and coarse, respectively, and the manual annotation of ground-truth is more feasible in 2d. The exact spatial and temporal location of these 2d navigator landmarks in the predicted 4d sequence is computed through linear interpolation. First, we validated our registration algorithm by comparing the landmark positions to their corresponding locations in the displacement fields. We computed the target registration error (TRE) for subject s over all n_s states of the annotated sequences (5 cycles) according to

$$\text{TRE}(s) = \frac{1}{2n_s} \sum_{i=1}^{n_s} \sum_{j=1}^2 |t_i(t_{s,j}^R) - l_{s,i,j}^T|, \quad (13)$$

where $l_{s,j}^R$ is the j -th landmark in the reference frame and $t_i(\cdot)$ a function that warps the landmark according to the registration result for frame i , and $l_{s,i,j}^T$ is the corresponding manually-selected landmark position. Table 3 gives the results (columns 2–7). The average registration error is within the same order as the in-plane resolution. Next, we used the landmarks instead of the registration results to compute the reconstruction error, see Table 3 (columns 8–13). These numbers are similar to Table 2. Again, Subject 4 shows the largest error. Fig. 7 shows that this error occurs at the inferior landmark near the inhalation states, whereas the curves match well for the superior landmark. Usually, the motion amplitude is lower in the inferior part of the liver compared to the superior part, but for Subject 4 it is the opposite. Also, in this case, the US points neither provide additional information about the inferior motion because they are located superiorly close to the diaphragm, which explains the worse results.

4.3. Evaluation of spatio-temporal prediction

In the spatio-temporal prediction experiments, a reconstruction of the organ position is computed for time $t + t_\Delta$ based on the temporally predicted US locations $\hat{\mathbf{p}}(t + t_\Delta)$ given by the ANN. Because temporal prediction introduces additional uncertainty, the regularisation factor η was increased to 3.0, determined empirically.

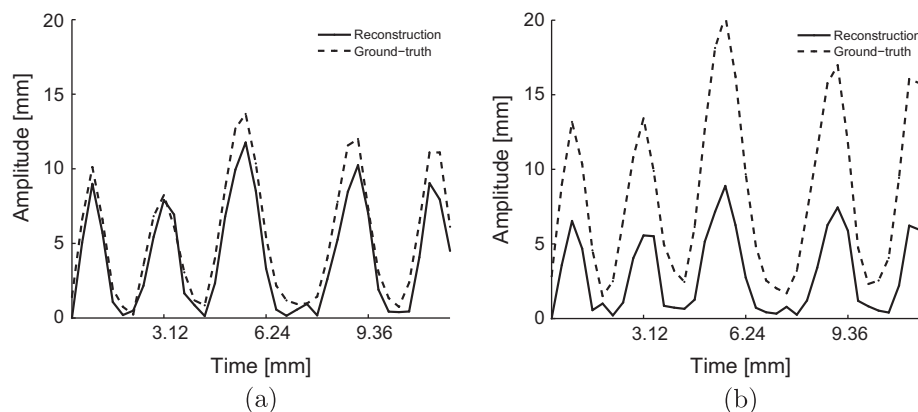


Fig. 7. Amplitude of reconstruction vs. manually identified ground-truth of observer 1 for subject 4. (a) Landmark centrally located on the navigator slice. (b) Landmark located near the inferior end of the liver.

Table 4
Results for spatio-temporal prediction for various prediction intervals t_Δ .

Subject	Mean error $e(t, i)$ (mm)								
	t_Δ (ms)	50	100	150	200	250	300	350	400
1		2.1	2.0	2.0	2.0	2.2	2.2	2.2	2.2
2		1.8	1.8	1.8	1.8	1.9	2.0	2.2	2.1
3		2.2	2.2	2.2	2.2	2.3	2.3	2.3	2.4
4		3.0	3.1	3.0	3.1	3.3	3.5	3.5	3.8
5		2.4	2.5	2.5	2.5	2.6	2.7	2.4	2.3
6		2.6	2.6	2.7	2.7	2.7	3.0	3.1	3.4
7		2.1	2.1	2.1	2.2	2.3	2.5	2.6	2.5
8		2.4	2.5	2.5	2.6	2.9	2.9	2.9	3.0
Mean		2.3	2.3	2.4	2.4	2.5	2.6	2.7	2.7

Table 4 summarises the results. The highest error is again observed for Subject 4. The mean error over all subjects ranges from 2.3 for 50 ms to 2.7 for 300 ms and increases monotonically.

The root mean square error (RMSE) is a common measure to compare a predicted signal at various times to its true value. Based on Eq. (12), we define it as follows:

$$\text{RMSE}(\tilde{\mathbf{x}}(t), \mathbf{x}(t)) = \sqrt{\frac{1}{nm} \sum_{t=1}^n \sum_{i=1}^m \|\tilde{\mathbf{x}}(t, i) - \mathbf{x}(t, i)\|^2}, \quad (14)$$

where n is the number of samples, *i.e.* length of the reconstructed sequence, and m the number of model vertices. Using the RMSE, we can compute the relative improvement achieved compared to spatial-prediction only. This measure is called the relative root mean square error (rRMSE) and is defined according to

$$\text{rRMSE}(\tilde{\mathbf{x}}(t), \mathbf{x}(t)) = \frac{\text{RMSE}(\tilde{\mathbf{x}}(t), \mathbf{x}(t))}{\text{RMSE}(\mathbf{x}(t - t_\Delta), \mathbf{x}(t))}, \quad (15)$$

where $\mathbf{x}(t - t_\Delta)$ is the ground-truth delayed according to the number of predicted frames. Fig. 8(a) shows a plot of the average RMSE with and without temporal prediction and Fig. 8(b) shows the average rRMSE over all subjects for the range of prediction times. With increasing t_Δ the relative improvement increases to almost 20%.

Fig. 9 shows slices through the mean exhalation shape with contours of the average error over all subjects for spatial prediction. The highest errors occur in the posterior-superior and inferior-anterior regions of the liver. A possible explanation for this is that the motion amplitude at locations near the diaphragm is highest and, additionally, both areas are farther away from the typical US plane position than most other points. As for the inferior part, the error is higher because of a higher variability in position and size of adjacent organs, *e.g.* kidney or intestine, leading to different motion behaviour among the subjects.

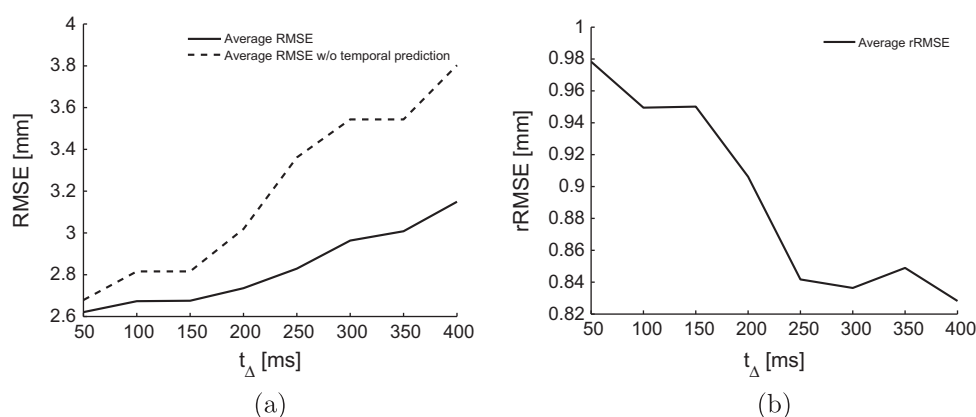


Fig. 8. (a) RMSE of spatio-temporal prediction vs. RMSE without temporal prediction for all values of t_{Δ} . (b) rRMSE curve for all values of t_{Δ} .

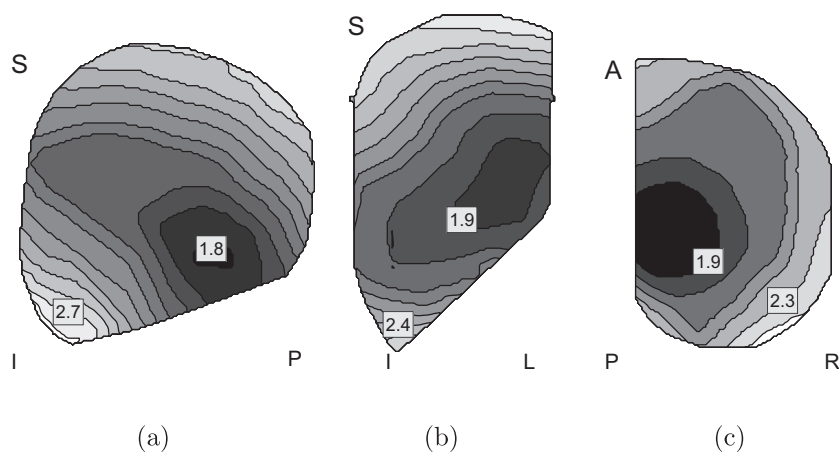


Fig. 9. (a) Sagittal, (b) coronal and (c) axial slice showing contour of the average error (in mm) for spatial prediction ($t_{\Delta} = 0$).

5. Conclusion and future work

We have presented a novel method to predict respiratory liver motion from points tracked in ultrasound (US) images both spatially and spatio-temporally. The novelty of this work is the combination of robust feature tracking in 2d US images for driving a population-based statistical motion model for spatial and spatio-temporal prediction and its validation using simultaneous 4D-MRI/US imaging. Prediction results are evaluated on an unprecedented amount of dense ground-truth motion extracted from 4D-MRI through deformable registration. Our approach is completely non-invasive and predicts the motion of the liver accurately over more than 5 m. Given a pre-operative scan of the liver for model initialisation, our statistical approach is solely driven by tracked points in 2d US images and does not make assumptions on either breathing depth or speed. Besides spatial-only prediction, we show how our model can be connected ahead with a temporal prediction algorithm in a straight-forward way. We show the efficiency of our spatio-temporal prediction approach by computing the rRMSE for a range of look-ahead times with encouraging results. Lastly, experiments on manually-selected landmarks confirm that the ground-truth motion obtained from deformable registration is accurate and that our prediction algorithm produces accurate results.

We envision the application of our technique in actual HIFU or proton therapy treatments and it remains to be shown how it performs on patients with liver tumours, as significantly different

motion patterns can be expected if a tumour is attached to surrounding structures. For HIFU, a typical treatment session presently takes several hours using state-of-the art technology, but the sonication is always fractioned into relative short portions. The duration of one fraction varies upon the application and technology, but the recent reports indicate a range between 4 s and 60 s (Petrusca et al., 2010; Salomir et al., 2013; Celicanin et al., 2013; Kim et al., 2011). This duration is significantly shorter than the temporal base of 5.5 m used as block acquisition and motion prediction in this study. Elementary HIFU lesion size including thermal diffusion enlarging effects is at least 3 mm in plane (Celicanin et al., 2013) and the clinically acceptable accuracy for motion prediction and compensation should therefore be less than 3 mm, which is achieved here. Finally, the approach described in this paper is not limited to the liver. Other upper abdominal organs such as pancreas or kidney can equally be motion-predicted using our framework.

Acknowledgement

The authors would like to thank the Swiss National Science Foundation for funding this project (CRSII2-127549).

References

- Amin, V., Wu, L., Long, T., Roberts, R., McClure, S., Ryken, T., 2008. Therapy planning and monitoring of tissue ablation by high intensity focused ultrasound (HIFU)

- using imaging and simulation. In: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4471.
- Arnold, P., Preiswerk, F., Fasel, B., Salomir, R., Scheffler, K., Cattin, P.C., 2011. 3D organ motion prediction for MR-guided high intensity focused ultrasound. In: Proceedings MICCAI, pp. 623–630.
- Balter, J.M., Brock, K.K., Litzenberg, D.W., McShan, D.L., Lawrence, T.S., Ten Haken, R., McGinn, C.J., Lam, K.L., Dawson, L.A., 2002. Daily targeting of intrahepatic tumors for radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 52, 266–271.
- Balter, J.M., Wright, J.N., Newell, L.J., Friemel, B., Dimmer, S., Cheng, Y., Wong, J., Vertatschitsch, E., Mate, T.P., 2005. Accuracy of a wireless localization system for radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 61, 933–937.
- Barber, C.B., Dobkin, D.P., Huhdanpaa, H., 1996. The quickhull algorithm for convex hulls. *ACM Trans. Math. Software* 22, 469–483.
- Blanz, V., Vetter, T., 2002. Reconstructing the complete 3D shape of faces from partial information. *Informationstechnik und Technische Informatik* 44, 295–302.
- Celicanin, Z., Auboiroux, V., Bieri, O., Petrusca, L., Santini, F., Viallon, M., Scheffler, K., Salomir, R., 2013. Real-time method for motion-compensated MR thermometry and MRgHIFU treatment in abdominal organs. *Magn. Reson. Med.*
- De Luca, V., Tanner, C., Székely, G., 2012. Speeding-up image registration for repetitive motion scenarios. In: Proceedings ISBI 2012. IEEE, pp. 1355–1358.
- Dryden, I., Mardia, K., 1998. *Statistical Shape Analysis*. John Wiley & Sons.
- Ehrhardt, J., Werner, R., Schmidt-Richberg, A., Handels, H., 2010. A statistical shape and motion model for the prediction of respiratory lung motion. In: Dawant, B.M., Haynor, D. (Eds.), *Image Processing, SPIE Medical Imaging 2010*, SPIE, San Diego, California, USA, pp. 762353-1–762353-9.
- Ehrhardt, J., Werner, R., Schmidt-Richberg, A., Handels, H., 2011. Statistical modeling of 4d respiratory lung motion using diffeomorphic image registration. *IEEE Trans. Med. Imag.* 30, 251–265.
- Ehrhardt, J., Werner, R., Schmidt-Richberg, A., Schulz, B., Handels, H., 2008. Generation of a mean motion model of the lung using 4D-CT image data. In: Proceedings of the First Eurographics conference on Visual Computing for Biomedicine. Eurographics Association, pp. 69–76.
- Eom, J., Xu, X.G., De, S., Shi, C., 2010. Predictive modeling of lung motion over the entire respiratory cycle using measured pressure-volume data, 4DCT images, and finite-element analysis. *Med. Phys.* 37, 4389–4400.
- Ernst, F., Driehen, R., Schlaefer, A., Schweikard, A., 2013. Evaluating and comparing algorithms for respiratory motion prediction. *Phys. Med. Biol.* 58, 3911.
- Fayad, H., Clement, J., Pan, T., Roux, C., Le Rest, C., Pradier, O., Visvikis, D., 2009. Towards a generic respiratory motion model for 4D CT imaging of the thorax. In: Nuclear Science Symposium Conference Record (NSS/MIC), 2009. IEEE, pp. 3975–3979.
- George, R., Chung, T.D., Vedam, S.S., Ramakrishnan, V., Mohan, R., Weiss, E., Keall, P.J., 2006. Audio-visual biofeedback for respiratory-gated radiotherapy: impact of audio instruction and audio-visual biofeedback on respiratory-gated radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 65, 924–933.
- Giebler, A., Fontenot, J., Balter, P., Ciangaru, G., Zhu, R., Newhauser, W., 2009. Dose perturbations from implanted helical gold markers in proton therapy of prostate cancer. *J. Appl. Clin. Med. Phys.* 10, 2875.
- Greco, C., Wolden, S., 2007. Current status of radiotherapy with proton and light ion beams. *Cancer* 109, 1227–1238.
- Hagan, M.T., Menhaj, M.B., 1994. Training feedforward networks with the marquardt algorithm. *IEEE Trans. Neural Networks* 5, 989–993.
- Hanley, J., Debois, M.M., Mah, D., Mageras, G.S., Raben, A., Rosenzweig, K., Mychalczak, B., Schwartz, L.H., Glogglger, P.J., Lutz, W., Ling, C.C., Leibel, S.A., Fuks, Z., Kutcher, G.J., 1999. Deep inspiration breath-hold technique for lung tumors: the potential value of target immobilization and reduced lung density in dose escalation. *Int. J. Radiat. Oncol. Biol. Phys.* 45, 603–611.
- He, T., Xue, Z., Xie, W., Wong, S.T.C., 2010. Online 4-D CT estimation for patient-specific respiratory motion based on real-time breathing signals. In: Proc. MICCAI 2010. Springer, pp. 392–399.
- Herfarth, K.K., Debus, J., Lohr, F., Bahner, M.L., Fritz, P., Hoss, A., Schlegel, W., Wannemacher, M.F., 2000. Extracranial stereotactic radiation therapy: set-up accuracy of patients treated for liver metastases. *Int. J. Radiat. Oncol. Biol. Phys.* 46, 329–335.
- Hoisak, J.D., Sixel, K.E., Tirona, R., Cheung, P.C., Pignol, J.P., 2006. Prediction of lung tumor position based on spirometry and on abdominal displacement: accuracy and reproducibility. *Radiother. Oncol.* 78, 339–346.
- Jiang, S.B., 2006. Technical aspects of image-guided respiration-gated radiation therapy. *Med. Dosim.* 31, 141–151.
- Keall, P.J., Kini, V.R., Vedam, S.S., Mohan, R., 2001. Motion adaptive X-ray therapy: a feasibility study. *Phys. Med. Biol.* 46, 1–10.
- Kim, Y.S., Lim, H.K., Kim, J.H., Rhim, H., Park, B.K., Keserci, B., Köhler, M.O., Bae, D.S., Kim, B.G., Lee, J.W., et al., 2011. Dynamic contrast-enhanced magnetic resonance imaging predicts immediate therapeutic response of magnetic resonance-guided high-intensity focused ultrasound ablation of symptomatic uterine fibroids. *Invest. Radiol.* 46, 639–647.
- King, A.P., Buerger, C., Tsoumpas, C., Marsden, P., Schaeffter, T., 2012. Thoracic respiratory motion estimation from mri using a statistical model and a 2-d image navigator. *Med. Image Anal.* 16, 252–264.
- Kiriyanthan, S., Fundana, K., Majeed, T., Cattin, P.C., 2012. A primal-dual approach for discontinuity preserving registration. In: 9th IEEE International Symposium on Biomedical Imaging (ISBI), pp. 350–353.
- Klinder, T., Lorenz, C., Ostermann, J., 2009. Free-breathing intra- and intersubject respiratory motion capturing, modeling, and prediction.
- Klinder, T., Lorenz, C., Ostermann, J., 2010. Prediction framework for statistical respiratory motion modeling. In: Jiang, T., Navab, N., Pluim, J., Viergever, M. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, Lecture Notes in Computer Science, vol. 6363. Springer, Berlin/Heidelberg, pp. 327–334.
- Koshani, R., Balter, J.M., Hayman, J.A., Henning, G.T., van Herk, M., 2006. Short-term and long-term reproducibility of lung tumor position using active breathing control (ABC). *Int. J. Radiat. Oncol. Biol. Phys.* 65, 1553–1559.
- Krauss, A., Nill, S., Oelfke, U., 2011. The comparative performance of four respiratory motion predictors for real-time tumour tracking. *Phys. Med. Biol.* 56, 5303.
- Kubo, H.D., Hill, B.C., 1996. Respiration gated radiotherapy treatment: a technical study. *Phys. Med. Biol.* 41, 83–91.
- Lomax, A., 1999. Intensity modulated methods for proton therapy. *Phys. Med. Biol.* 44, 185–205.
- Lu, W., Ruchala, K.J., Chen, M.L., Chen, Q., Olivera, G.H., 2006. Real-time respiration monitoring using the radiotherapy treatment beam. *Phys. Med. Biol.* 51, 4469–4495.
- McClelland, J., Hawkes, D.J., Schaeffter, T., King, A.P., 2013. Respiratory motion models: a review. *Med. Image Anal.* 17, 19–42.
- Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S., 2010. Fast free-form deformation using graphics processing units. *Comput. Methods Prog. Biomed.* 98, 278–284.
- Murphy, M.J., 2004. Tracking moving organs in real time. *Semin. Radiat. Oncol.* 14, 91–100.
- Nelson, C., Starkschall, G., Balter, P., Fitzpatrick, M.J., Antolak, J.A., Tolani, N., Prado, K., 2005. Respiration-correlated treatment delivery using feedback-guided breath hold: a technical study. *Med. Phys.* 32, 175–181.
- Ohara, K., Okumura, T., Akisada, M., Inada, T., Mori, T., Yokota, H., Calaguas, M.J.B., 1989. Irradiation synchronized with respiration gate. *Int. J. Radiat. Oncol. Biol. Phys.* 17, 853–857.
- Palussiere, J., Salomir, R., Le Bail, B., Fawaz, R., Quesson, B., Grenier, N., Moonen, C.T.W., 2003. Feasibility of MR-guided focused ultrasound with real-time temperature mapping and continuous sonication for ablation of VX2 carcinoma in rabbit thigh. *Magn. Reson. Med.* 49, 89–98.
- Pedersen, A.N., Korreman, S., Nystrom, H., Specht, L., 2004. Breathing adapted radiotherapy of breast cancer: reduction of cardiac and pulmonary doses using voluntary inspiration breath-hold. *Radiother. Oncol.* 72, 53–60.
- Petrusca, L., Cattin, P., De Luca, V., Preiswerk, F., Celicanin, Z., Auboiroux, V., Viallon, M., Arnold, P., Santini, F., Terraz, S., Scheffler, K., Becker, C.D., Salomir, R., 2013. Hybrid ultrasound/magnetic resonance simultaneous acquisition and image fusion for motion monitoring in the upper abdomen. *Invest. Radiol.* 48, 333–340.
- Petrusca, L., Salomir, R., Brassat, L., Chavrier, F., Cotton, F., Chapelon, J.Y., 2010. Sector-switching sonication strategy for accelerated HIFU treatment of prostate cancer: in vitro experimental validation. *IEEE Trans. Biomed. Eng.* 57, 17–23.
- Preiswerk, F., Arnold, P., Fasel, B., Cattin, P., 2012a. A bayesian framework for estimating respiratory liver motion from sparse measurements. In: Yoshida, H., Sakas, G., Linguraru, M. (Eds.), *Abdominal Imaging, Computational and Clinical Applications, Lecture Notes in Computer Science*, vol. 7029. Springer, Berlin/Heidelberg, pp. 207–214.
- Preiswerk, F., Arnold, P., Fasel, B., Cattin, P.C., 2012b. Robust tumour tracking from 2d imaging using a population-based statistical motion model. In: IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA), 2012, pp. 209–214.
- Putra, D., Haas, O.C.L., Mills, J.A., Burnham, K.J., 2008. A multiple model approach to respiratory motion prediction for real-time IGRT. *Phys. Med. Biol.* 53, 1651.
- Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imag.* 18, 712–721.
- Salomir, R., Petrusca, L., Auboiroux, V., Muller, A., Vargas, M.I., Morel, D.R., Goget, T., Breguet, R., Terraz, S., Hopple, J., Montet, X., Becker, C.D., Viallon, M., 2013. Magnetic resonance-guided shielding of prefocal acoustic obstacles in focused ultrasound therapy: application to intercostal ablation in liver. *Invest. Radiol.* 48, 366–380.
- Seppenwoolde, Y., Berbeco, R.I., Nishioka, S., Shirato, H., Heijmen, B., 2007. Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study. *Med. Phys.* 34, 2774–2784.
- Sharp, G.C., Jiang, S.B., Shimizu, S., Shirato, H., 2004. Prediction of respiratory tumour motion for real-time image-guided radiotherapy. *Phys. Med. Biol.* 49, 425–440.
- Shirato, H., Seppenwoolde, Y., Kitamura, K., Onimura, R., Shimizu, S., 2004. Intrafractional tumor motion: lung and liver. *Semin. Radiat. Oncol.* 14, 10–18.
- Simon, L., Giraud, P., Serois, V., Rosenwald, J.C., 2005. Lung volume assessment for a cross-comparison of two breathing-adapted techniques in radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 63, 602–609.
- Sundaram, T., Avants, B., Gee, J., 2004. A dynamic model of average lung deformation using capacity-based reparameterization and shape averaging of lung mr images. In: Barillot, C., Haynor, D., Hellier, P. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004*, Lecture Notes in Computer Science, vol. 3217. Springer, Berlin/Heidelberg, pp. 1000–1007.
- Tanner, C., Boye, D., Samei, G., Székely, G., 2012. Review on 4d models for organ motion compensation. *Crit. Rev. Biomed. Eng.* 40, 135–154.

- Uematsu, M., Shioda, A., Suda, A., Tahara, K., Kojima, T., Hama, Y., Kono, M., Wong, J.R., Fukui, T., Kusano, S., 2000. Intrafractional tumor position stability during computed tomography (CT)-guided frameless stereotactic radiation therapy for lung or liver cancers with a fusion of CT and linear accelerator (FOCAL) unit. *Int. J. Radiat. Oncol. Biol. Phys.* 48, 443–448.
- Vedam, S.S., Kini, V.R., Keall, P.J., Ramakrishnan, V., Mostafavi, H., Mohan, R., 2003. Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker. *Med. Phys.* 30, 505–513.
- von Siebenthal, M., Székely, G., Gamper, U., Boesiger, P., Lomax, A., Cattin, P., 2007a. 4D MR imaging of respiratory organ motion and its variability. *Phys. Med. Biol.* 52, 1547–1564.
- von Siebenthal, M., Székely, G., Lomax, A., Cattin, P., 2007b. Systematic errors in respiratory gating due to intrafraction deformations of the liver. *Med. Phys.* 34, 3620–3629.
- Welch, G., Bishop, G., 1995. An introduction to the Kalman filter.
- Werner, R., Ehrhardt, J., Schmidt, R., Handels, H., 2009. Patient-specific finite element modeling of respiratory lung motion using 4D CT image data. *Med. Phys.* 36, 1500–1511.
- Zhang, T., Orton, N.P., Tome, W.A., 2005. On the automated definition of mobile target volumes from 4D-CT images for stereotactic body radiotherapy. *Med. Phys.* 32, 3493–3502.

Chapter 8

Conclusions

In this thesis, a population-based statistical model for respiratory organ motion together with an algorithm for spatial prediction of the complete organ from sparse measurements is proposed. During prediction, sparse input of the liver is obtained to predict the position of the complete organ. Besides various simulations with 3d, 2d and 1d displacements of the liver, the combination of the model with tracked points in 2d ultrasound is particularly studied in a clinical setup. The use of ultrasound as a cheap non-invasive modality turns out to be well suited for this purpose. The average error over long prediction session is between 2 mm and 3 mm for most subjects. We also show how the prediction algorithm can be combined with any temporal predictor to perform spatio-temporal motion prediction and the results are in the same range. These results, together with results from previous simulations, show that our technique is indeed interesting for motion compensation in tumour therapy.

The approach described in this thesis is advantageous in various aspects. First, no subject-specific motion data must be acquired to use the method on a new subject. Today's techniques for 3d imaging are either still cost-intensive (MRI) or expose the subject to significant amounts of radiation (CT). Neither of the two is required in our approach, except for a pre-operative 3d scan (which is standard procedure for dose calculation in radiotherapy) to align the model to the subject's anatomy. The method is therefore economical with respect to labour, cost and only requires to expose the patient to a minimal amount of radiation, if at all. During prediction, we also do not rely on any invasive technology, as

the respiratory signal is obtained from ultrasound, although the model can also be used for prediction based on implanted electromagnetic beacons or any other method that provides fiducial information of the liver.

The ultimate goal in motion compensation is to continuously irradiate a moving tumour. We hope that our contribution is a step in this direction. However, we also envision interesting applications in less rigorous scenarios, *e.g.* respiratory gating, as our model is more precise in the vicinity of exhalation. It can be used to push the limits of traditional respiratory gating by compensating for respiratory motion only within the gating window. The use of ultrasound to obtain a respiratory signal for this purpose directly from the organ is a significant improvement on its own compared to traditional gating, where the signal is often obtained from a spirometer or displacements of the chest wall. Taking advantage of model based spatio-temporal prediction on top of this can be a next step in tumour therapy and pave the way to more precise and cheaper tumour treatment. At the time of writing, the ideas developed in this work have been picked up by other researchers in the field [35, 36, 74].

There are numerous areas for improvement and future work. For model building it would be helpful to have a better understanding on how accurate 4D-MRI is and how much data is necessary to obtain 3d reconstructions of sufficient quality. Given this knowledge, much effort can be invested in optimising the 4D-MRI sequence to further push image quality and temporal resolution. A thorough evaluation of the employed registration method and its limitations may help to understand how some of the prediction error arises and to separate error due to inaccuracies in the ground-truth from actual mispredictions. A known limitation of virtually all of today's established registration algorithms is their disability to describe discontinuous motion. There are interesting approaches which tackle exactly this problem [75, 76] and, although still computationally expensive, they might one day provide efficient registration for discontinuous motion. Regarding the acquisition of fiducial information, extending to 2.5d or 3d ultrasound is an interesting option to obtain a larger field of view and improved depth control of the tracked points. Introducing further knowledge about the respiratory phase to the reconstruction algorithm, *e.g.* using a temporal regularisation term, could be a way to further improve the stability of spatial predictions.

References

- [1] P. Boyle, B. Levin, and I. A. for Research on Cancer, *World cancer report 2008*. IARC Nonserial Publication Series, IARC Press, 2008. [1](#)
- [2] G. Delaney, M. Barton, and S. Jacob, “Estimation of an optimal radiotherapy utilization rate for breast carcinoma,” *Cancer*, vol. 98, no. 9, pp. 1977–1986, 2003. [1](#)
- [3] C. Yu, D. Jaffray, and J. Wong, “The effects of intra-fraction organ motion on the delivery of dynamic intensity modulation,” *Phys. Med. Biol.*, vol. 43, no. 1, pp. 91–104, 1998. [1](#)
- [4] P. J. Keall, V. R. Kini, S. S. Vedam, and R. Mohan, “Motion adaptive x-ray therapy: a feasibility study,” *Phys. Med. Biol.*, vol. 46, no. 1, pp. 1–10, 2001. [1](#), [9](#)
- [5] J. Mechalakos, E. Yorke, G. S. Mageras, A. Hertanto, A. Jackson, C. Obcemea, K. Rosenzweig, and C. C. Ling, “Dosimetric effect of respiratory motion in external beam radiotherapy of the lung,” *Radiother. Oncol.*, vol. 71, no. 2, pp. 191–200, 2004. [1](#)
- [6] M. von Siebenthal, G. Székely, A. Lomax, and P. Cattin, “Systematic errors in respiratory gating due to intrafraction deformations of the liver,” *Med. Phys.*, vol. 34, no. 9, pp. 3620–3629, 2007. [1](#), [9](#), [10](#)
- [7] A. Lomax, “Intensity modulated methods for proton therapy,” *Phys. Med. Biol.*, vol. 44, pp. 185–205, 1999. [2](#), [7](#)

REFERENCES

- [8] V. Amin, L. Wu, T. Long, R. Roberts, S. McClure, and T. Ryken, "Therapy planning and monitoring of tissue ablation by high intensity focused ultrasound (hifu) using imaging and simulation," in *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, p. 4471, 2008. [2](#), [7](#)
- [9] M. A. Penna, K. A. Dines, R. Seip, R. F. Carlson, and N. T. Sanghvi, "Modeling prostate anatomy from multiple view TRUS images for image-guided HIFU therapy," *IEEE Transactions On Ultrasonics Ferroelectrics And Frequency Control*, vol. 54, pp. 52–69, JAN 2007. [2](#), [7](#)
- [10] J. Palussiere, R. Salomir, B. Le Bail, R. Fawaz, B. Quesson, N. Grenier, and C. T. W. Moonen, "Feasibility of MR-guided focused ultrasound with real-time temperature mapping and continuous sonication for ablation of VX2 carcinoma in rabbit thigh," *Magnetic Resonance In Medicine*, vol. 49, pp. 89–98, JAN 2003. [2](#), [7](#)
- [11] 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.*, vol. 52, no. 6, pp. 1547–1564, 2007. [2](#), [10](#)
- [12] J. B. West, *Respiratory Physiology: The Essentials*. Lippincott Williams & Wilkins, ninth edition ed., 2011. [3](#)
- [13] 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, C. R. Ramsey, M. B. van Herk, S. S. Vedam, J. W. Wong, and E. Yorke, "The management of respiratory motion in radiation oncology report of AAPM task group 76," *Med. Phys.*, vol. 33, no. 10, pp. 3874–3900, 2006. [3](#), [4](#), [8](#)
- [14] S. C. Davies, A. L. Hill, R. B. Holmes, M. Halliwell, and P. C. Jackson, "Ultrasound quantitation of respiratory organ motion in the upper abdomen," *Brit. J. Radiol.*, vol. 67, no. 803, pp. 1096–1102, 1994. [4](#)
- [15] H. Korin, R. Ehman, S. Riederer, J. Felmlee, and R. Grimm, "Respiratory kinematics of the upper abdominal organs - a quantitative study," *Magnetic Resonance In Medicine*, vol. 23, no. 1, pp. 172–178, 1992. [4](#)

REFERENCES

- [16] M. von Siebenthal, *Analysis and Modelling of Respiratory Liver Motion using 4DMRI*. PhD thesis, ETH Zurich, 2008. No. 17613. [4](#), [11](#), [13](#), [15](#), [17](#), [22](#)
- [17] I. Suramo, M. Paivansalo, and V. Myllyla, “Cranio-caudal movements of the liver, pancreas and kidneys in respiration,” *Acta Radiol. Diagn.*, vol. 25, pp. 129–131, 1984. [4](#)
- [18] P. J. Bryan, S. Custar, J. R. Haaga, and V. Balsara, “Respiratory movement of the pancreas: an ultrasonic study,” *Journal of ultrasound in medicine*, vol. 3, pp. 317–320, July 1984. [4](#)
- [19] P. H. Weiss, J. M. Baker, and E. J. Potchen, “Assessment of hepatic respiratory excursion,” *J Nucl Med*, vol. 13, no. 10, pp. 758–9, 1972. [4](#)
- [20] G. Harauz and M. J. Bronskill, “Comparison of the liver’s respiratory motion in the supine and upright positions: concise communication,” *J Nucl Med*, vol. 20, no. 7, pp. 733–5, 1979. [4](#)
- [21] O. Wade, “Movements of the thoracic cage and diaphragm in respiration,” *J Physiol*, vol. 124, no. 2, pp. 758–9, 1954. [4](#)
- [22] P. Giraud, Y. D. Rycke, B. Dubray, S. Helfre, D. Voican, L. Guo, J. Rosenwald, K. Keraudy, M. Housset, E. Touboul, and J. Cosset, “Conformal radiotherapy (crt) planning for lung cancer: Analysis of intrathoracic organ motion during extreme phases of breathing,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 51, no. 4, pp. 1081–1092, 2001. [4](#)
- [23] E. C. Ford, G. S. Mageras, E. Yorke, K. E. Rosenzweig, R. Wagman, and C. C. Ling, “Evaluation of respiratory movement during gated radiotherapy using film and electronic portal imaging,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 52, no. 2, pp. 522–531, 2002. [4](#)
- [24] Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. van Herk, J. 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. Biol. Phys.*, vol. 53, no. 4, pp. 822–834, 2002. [5](#)

REFERENCES

- [25] M. Chaudry and M. Winslet, *Surgical Oncology*. Oxford Specialist Handbooks in Surgery, OUP Oxford, 2009. 5
- [26] R. Skeel and S. Khleif, *Handbook of Cancer Chemotherapy*. Lippincott Williams & Wilkins, 2011. 5
- [27] J. P. Mcgahan and G. D. Dodd, “Radiofrequency Ablation of the Liver: Current Status,” *Am. J. Roentgenol.*, vol. 176, pp. 3–16, Jan. 2001. 5
- [28] L. M. Sutherland, J. A. R. Williams, R. T. A. Padbury, D. C. Gotley, B. Stokes, and G. J. Maddern, “Radiofrequency ablation of liver tumors: a systematic review,” *Arch Surg*, vol. 141, no. 2, pp. 181–90, 2006. 5
- [29] J. M. Llovet, R. Vilana, C. Brú, L. Bianchi, J. M. Salmeron, L. Boix, S. Ganau, M. Sala, M. Pagès, C. Ayuso, M. Solé, J. Rodés, and J. Bruix, “Increased risk of tumor seeding after percutaneous radiofrequency ablation for single hepatocellular carcinoma,” *Hepatology*, vol. 33, no. 5, pp. 1124–9, 2001. 6
- [30] J. Galvin, G. Ezzell, A. Eisbrauch, C. Yu, B. Butler, Y. Xiao, I. Rosen, J. Rosenman, M. Sharpe, and L. Xing, “Implementing IMRT in clinical practice: a joint document of the American Society for Therapeutic Radiology and Oncology and the American Association of Physicists in Medicine,” *International Journal of Radiation Oncology Biology Physics*, vol. 58, pp. 1616–1634, Apr. 2004. 6
- [31] N. M. Aziz and J. H. Rowland, “Trends and advances in cancer survivorship research: challenge and opportunity,” *Semin Radiat Oncol*, vol. 13, pp. 248–266, July 2003. 6
- [32] R. Curtis, *New Malignancies Among Cancer Survivors: SEER Cancer Registries, 1973-2000*. NIH publication, U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute, 2006. 6
- [33] R. F. Barth, M. G. H. Vicente, O. K. Harling, W. S. Kiger, K. J. Riley, P. J. Binns, F. M. Wagner, M. Suzuki, T. Aihara, I. Kato, and S. Kawabata,

REFERENCES

- “Current status of boron neutron capture therapy of high grade gliomas and recurrent head and neck cancer,” *Radiat Oncol*, vol. 7, p. 146, 2012. 7
- [34] D. Ash, A. Gerbaulet, E. S. for Therapeutic Radiology, and Oncology, *The GEC ESTRO Handbook of Brachytherapy*. European Society for Therapeutic Radiology and Oncology, 2002. 7
- [35] J. McClelland, D. J. Hawkes, T. Schaeffter, and A. P. King, “Respiratory motion models: A review,” *Medical Image Analysis*, vol. 17, no. 1, pp. 19–42, 2013. 8, 72
- [36] C. Tanner, D. Boye, G. Samei, and G. Szekely, “Review on 4d models for organ motion compensation,” *Critical Reviews in Biomedical Engineering*, vol. 40, no. 2, pp. 135–154, 2012. 8, 72
- [37] A. N. Pedersen, S. Korreman, H. Nystrom, and L. Specht, “Breathing adapted radiotherapy of breast cancer: reduction of cardiac and pulmonary doses using voluntary inspiration breath-hold,” *Radiother. Oncol.*, vol. 72, no. 1, pp. 53–60, 2004. 9
- [38] C. Nelson, G. Starkschall, P. Balter, M. J. Fitzpatrick, J. A. Antolak, N. Tolani, and K. Prado, “Respiration-correlated treatment delivery using feedback-guided breath hold: A technical study,” *Med. Phys.*, vol. 32, no. 1, pp. 175–181, 2005. 9
- [39] J. Hanley, M. M. Debois, D. Mah, G. S. Mageras, A. Raben, K. Rosenzweig, B. Mychalczak, L. H. Schwartz, P. J. Gloeggler, W. Lutz, C. C. Ling, S. A. Leibel, Z. Fuks, and G. J. Kutcher, “Deep inspiration breath-hold technique for lung tumors: the potential value of target immobilization and reduced lung density in dose escalation,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 45, no. 3, pp. 603–611, 1999. 9
- [40] J. M. Balter, K. K. Brock, D. W. Litzenberg, D. L. McShan, T. S. Lawrence, R. Ten Haken, C. J. McGinn, K. L. Lam, and L. A. Dawson, “Daily targeting of intrahepatic tumors for radiotherapy,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 52, no. 1, pp. 266–271, 2002. 9

REFERENCES

- [41] R. Koshani, J. M. Balter, J. A. Hayman, G. T. Henning, and M. van Herk, “Short-term and long-term reproducibility of lung tumor position using active breathing control (ABC),” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 65, no. 5, pp. 1553–1559, 2006. [9](#)
- [42] M. Uematsu, A. Shioda, A. Suda, K. Tahara, T. Kojima, Y. Hama, M. Kono, J. R. Wong, T. Fukui, and S. Kusano, “Intrafractional tumor position stability during computed tomography (CT)-guided frameless stereotactic radiation therapy for lung or liver cancers with a fusion of CT and linear accelerator (FOCAL) unit,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 48, no. 2, pp. 443–448, 2000. [9](#)
- [43] T. Zhang, N. P. Orton, and W. A. Tome, “On the automated definition of mobile target volumes from 4D-CT images for stereotactic body radiotherapy,” *Med. Phys.*, vol. 32, no. 11, pp. 3493–3502, 2005. [9](#)
- [44] K. K. Herfarth, J. Debus, F. Lohr, M. L. Bahner, P. Fritz, A. Hoss, W. Schlegel, and M. F. Wannemacher, “Extracranial stereotactic radiation therapy: set-up accuracy of patients treated for liver metastases,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 46, no. 2, pp. 329–335, 2000. [9](#)
- [45] W. Lu, K. J. Ruchala, M.-L. Chen, Q. Chen, and G. H. Olivera, “Real-time respiration monitoring using the radiotherapy treatment beam,” *Phys. Med. Biol.*, vol. 51, no. 18, pp. 4469–4495, 2006. [9](#)
- [46] 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. Biol. Phys.*, vol. 17, no. 4, pp. 853–857, 1989. [9](#)
- [47] H. D. Kubo and B. C. Hill, “Respiration gated radiotherapy treatment: a technical study,” *Phys. Med. Biol.*, vol. 41, no. 1, pp. 83–91, 1996. [9](#)
- [48] R. George, T. D. Chung, S. S. Vedam, V. Ramakrishnan, R. Mohan, E. Weiss, and P. J. Keall, “Audio-visual biofeedback for respiratory-gated radiotherapy: Impact of audio instruction and audio-visual biofeedback on respiratory-gated radiotherapy,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 65, no. 3, pp. 924–933, 2006. [9](#)

REFERENCES

- [49] S. B. Jiang, “Technical aspects of image-guided respiration-gated radiation therapy,” *Medical Dosimetry*, vol. 31, no. 2, pp. 141–151, 2006. [9](#)
- [50] H. Shirato, S. Shimizu, K. Kitamura, T. Nishioka, K. Kagei, S. Hashimoto, H. Aoyama, T. Kunieda, N. Shinohara, H. Dosaka-Akita, and K. Miyasaka, “Four-dimensional treatment planning and fluoroscopic real-time tumor tracking radiotherapy for moving tumor,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 48, no. 2, pp. 435–442, 2000. [9](#)
- [51] G. C. Sharp, S. B. Jiang, S. Shimizu, and H. Shirato, “Prediction of respiratory tumour motion for real-time image-guided radiotherapy,” *Phys. Med. Biol.*, vol. 49, no. 3, pp. 425–440, 2004. [9](#)
- [52] S. S. Vedam, P. J. Keall, A. Docef, D. A. Todor, V. R. Kini, and R. Mohan, “Predicting respiratory motion for four-dimensional radiotherapy,” *Med. Phys.*, vol. 31, no. 8, pp. 2274–2283, 2004. [9](#)
- [53] L. Simon, P. Giraud, V. Servois, and J.-C. Rosenwald, “Lung volume assessment for a cross-comparison of two breathing-adapted techniques in radiotherapy,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 63, no. 2, pp. 602–609, 2005. [9](#), [12](#)
- [54] M. J. Murphy, “Tracking moving organs in real time,” *Semin. Radiat. Oncol.*, vol. 14, no. 1, pp. 91–100, 2004. [9](#)
- [55] S. S. Vedam, V. R. Kini, P. J. Keall, V. Ramakrishnan, H. Mostafavi, and R. Mohan, “Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker,” *Med. Phys.*, vol. 30, no. 4, pp. 505–513, 2003. [9](#)
- [56] J. D. Hoisak, K. E. Sixel, R. Tirona, P. C. Cheung, and J.-P. Pignol, “Prediction of lung tumour position based on spirometry and on abdominal displacement: Accuracy and reproducibility,” *Radiother. Oncol.*, vol. 78, no. 3, pp. 339–346, 2006. [9](#)

REFERENCES

- [57] R. D. Wiersma, W. Mao, and L. Xing, “Combined kV and MV imaging for real-time tracking of implanted fiducial markers,” *Medical Physics*, vol. 35, p. 1191, 2008. [10](#), [12](#)
- [58] H. Shirato, S. Shimizu, T. Shimizu, T. Nishioka, and K. Miyasaka, “Real-time tumour-tracking radiotherapy,” *The Lancet*, vol. 353, no. 9161, pp. 1331 – 1332, 1999. [12](#)
- [59] P. Kupelian, T. Willoughby, A. Mahadevan, T. Djemil, G. Weinstein, S. Jani, C. Enke, T. Solberg, N. Flores, D. Liu, D. Beyer, and L. Levine, “Multi-institutional clinical experience with the calypso system in localization and continuous, real-time monitoring of the prostate gland during external radiotherapy,” *International Journal of Radiation Oncology, Biology, Physics*, vol. 67, no. 4, pp. 1088 – 1098, 2007. [12](#)
- [60] P. Keall, A. Todor, S. Vedam, C. Barteel, J. Siebers, V. Kini, and R. Mohan, “On the use of EPID-based implanted marker tracking for 4D radiotherapy,” *Med. Phys.*, vol. 31, no. 12, pp. 3492–3499, 2004. [12](#)
- [61] J. P. Thirion, “Image matching as a diffusion process: an analogy with maxwell’s demons,” *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, 1998. [15](#)
- [62] D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D. Hawkes, “Non-rigid registration of breast MR images using mutual information,” in *MICCAI 1998*, pp. 1144–1152, 1998. [15](#)
- [63] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, “Automated multi-modality image registration based on information theory,” in *Proc. IPMI*, pp. 263–274, 1995. [15](#)
- [64] P. Viola and W. M. Wells, “Alignment by maximization of mutual information,” in *Proc. ICCV ’95*, pp. 16–23, IEEE Computer Society, 1995. [15](#)
- [65] T. F. Cootes, D. H. Cooper, C. J. Taylor, and J. Graham, “Trainable method of parametric shape-description,” *Image Vision Comput.*, vol. 10, no. 5, pp. 289–294, 1992. [17](#)

REFERENCES

- [66] T. Cootes, C. Taylor, D. H. Cooper, and J. Graham, “Training models of shape from sets of examples,” in *In Proc. British Machine Vision Conference*, pp. 9–18, 1992. [17](#)
- [67] V. Blanz and T. Vetter, “A morphable model for the synthesis of 3d faces,” in *Proceedings of SIGGRAPH*, pp. 187–194, 1999. [17](#)
- [68] Z. Tu, “Learning Generative Models via Discriminative Approaches,” in *Computer Vision and Pattern Recognition, 2007. IEEE Conference on*, pp. 1–8, IEEE, June 2007. [18](#)
- [69] R. Knothe, *A Global-to-Local Model for the Representation of Human Faces*. PhD thesis, University of Basel, Switzerland, 2009. [18](#)
- [70] I. Dryden and K. Mardia, *Statistical Shape Analysis*. John Wiley & Sons, 1998. [23](#)
- [71] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, “The quickhull algorithm for convex hulls,” *ACM T. Math. Software*, vol. 22, no. 4, pp. 469–483, 1996. [24](#)
- [72] M. von Siebenthal, G. Székely, A. Lomax, and P. Cattin, “Inter-subject modelling of liver deformation during radiation therapy,” in *MICCAI 2007*, vol. 4791 of *LNCS*, pp. 659–666, Springer, 2007. [24](#)
- [73] V. Blanz and T. Vetter, “Reconstructing the complete 3D shape of faces from partial information,” *Informationstechnik und Technische Informatik*, vol. 44, no. 6, pp. 295–302, 2002. [29](#), [32](#)
- [74] G. Samei, C. Tanner, and G. Székely, “Predicting liver motion using exemplar models,” in *Abdominal Imaging*, pp. 147–157, 2012. [72](#)
- [75] S. Kiriyanthan, K. Fundana, and P. C. Cattin, “Discontinuity Preserving Registration of Abdominal MR Images with Apparent Sliding Organ Motion,” in *Abdominal Imaging. Computational and Clinical Applications*, vol. 7029 of *LNCS*, pp. 231–139, Springer-Verlag, 2012. [72](#)

REFERENCES

- [76] S. Kiriyanthan, K. Fundana, T. Majeed, and P. C. Cattin, “A primal-dual approach for discontinuity preserving registration,” in *9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 350–353, 2012. [72](#)

Curriculum Vitae

Personal Data

Name	Frank Preiswerk
Date of birth	27. March 1982
Citizenship	Basel, Switzerland

Education

2010–2013	PhD in Biomedical Engineering Medical Image Analysis Center, University of Basel, Switzerland Supervisor: Prof. Dr. Philippe C. Cattin
2006–2008	M.Sc. in Computer Science Department of Computer Science, University of Basel, Switzerland
2003–2006	B.Sc. in Computer Science Department of Computer Science, University of Basel, Switzerland

References

Prof. Dr. Philippe C. Cattin
Medical Image Analysis Center (MIAC)
University of Basel, Switzerland
+41 (0)61 265 96 55
philippe.cattin@unibas.ch

Prof. Dr. Thomas Vetter
Department of Mathematics and Computer Science
University of Basel, Switzerland
+41 (0)61 267 05 56
thomas.vetter@unibas.ch

Publications

- 2014 Frank Preiswerk, Valeria De Luca, Patrik Arnold, Zarko Celicanin, Lorena Petrusca, Christine Tanner, Oliver Bieri, Rares Salomir, and Philippe C. Cattin. Model-guided respiratory organ motion prediction of the liver from 2d ultrasound. *Medical Image Analysis*, 18(5):740 – 751, 2014
- Lorena Petrusca, Philippe Cattin, Valeria De Luca, Frank Preiswerk, Zarko Celicanin, Vincent Auboiroux, Magalie Viallon, Patrik Arnold, Francesco Santini, Sylvain Terraz, Klaus Scheffler, Christoph D. Becker, , and Rares Salomir. Hybrid ultrasound/magnetic resonance simultaneous acquisition and image fusion for motion monitoring in the upper abdomen. *Investigative Radiology*, 48(5):333–340, May 2013
- 2012 Frank Preiswerk, Patrik Arnold, Beat Fasel, and Philippe Cattin. A bayesian framework for estimating respiratory liver motion from sparse measurements. In Hiroyuki Yoshida, Georgios Sakas, and Marius Linguraru, editors, *Abdominal Imaging. Computational and Clinical Applications*, volume 7029 of *Lecture Notes in Computer Science*, pages 207–214. Springer Berlin / Heidelberg, 2012
- Frank Preiswerk, Patrik Arnold, Beat Fasel, and Philippe C. Cattin. Robust tumour tracking from 2d imaging using a population-based statistical motion model. In *Mathematical Methods in Biomedical Image Analysis (MMBIA), 2012 IEEE Workshop on*, pages 209 –214, jan. 2012
- Frank Preiswerk, Patrik Arnold, Beat Fasel, and Philippe C. Cattin. Towards more precise, minimally-invasive tumour treatment under free breathing. *34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE EMBC 2012*, 9 2012
- Patrik Arnold, Frank Preiswerk, Beat Fasel, Rares Salomir, Klaus Scheffler, and Philippe Cattin. Model-based respiratory motion compensation in mrghifu. *IPCAI Proceedings*, 2012
- 2011 Patrik Arnold, Frank Preiswerk, Beat Fasel, Rares Salomir, Klaus Scheffler, and Philippe Cattin. 3d organ motion prediction for mr-guided high intensity focused ultrasound. In *Medical Image Computing and Computer-Assisted Intervention MICCAI 2011*, volume 6892 of *Lecture Notes in Computer Science*, pages 623–630. Springer Berlin / Heidelberg, 2011
- Z. Celicanin, F. Preiswerk, P. Arnold, P. Cattin, K. Scheffler, and F. Santini. Simultaneous acquisition of image and navigator slices using caipirinha. In *ISMRM*, 2011

Talks

- 2012 Paper presentation at the *Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*, January 2012, Breckenridge CO, USA
- 2011 Paper presentation at the *Workshop on Computational and Clinical Applications in Abdominal Imaging (ABD-MICCAI)*, September 2011, Toronto, Canada
- Invited talk at *Accuray Inc*, September 2011, Sunnyvale CA, USA

Other

- 2012 Member of the Program- and Scientific Review Commmittees of the *Workshop on Computational and Clinical Applications in Abdominal Imaging (ABD-MICCAI)*, October 2012, Nice, France