

Respiratory-Induced Organ Motion Compensation for MRgHIFU

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Medizinischen Fakultät
der Universität Basel

von

Patrik Arnold
aus Gümligen, Bern

Basel, 2012

Genehmigt von der Medizinischen Fakultät
auf Antrag von

Prof. Dr. Philippe C. Cattin, Basel
Dissertationsbetreuer

Prof. Dr. Antony Lomax, PSI, Villigen
Koreferent

Prof. Dr. med. Frank Zimmermann, Basel
Externer Gutachter

Basel, den

Prof. Dr. med. Christoph Beglinger
Dekan

Preface

Medical Image Analysis is an exciting and emerging field of research. The combination of different academic disciplines such as computer science, medicine, mathematics and physics to one interdisciplinary research field involves great challenges and opportunities at the same time.

With a background in applied physics and an area of specialisation in atmospheric microwave physics, I was rather new in the discipline. Getting into this field and finally writing this thesis would have been impossible without the help and dedicated guidance of many people.

First and foremost, I would like to thank my supervisor Prof. Dr. Philippe C. Cattin. I thank Philippe for his superb guidance, his positivism and his infectious enthusiasm in a wide range of topics. Being part of his group was always a great pleasure. Throughout the years, I have had the pleasure of working with many funny and inspiring colleagues in Basel. In particular, I would like to thank my fellow PhD student Frank Preiswerk for the close scientific collaboration and for designing almost everything I have ever published. I thank Beat Fasel for the profound GNU/Linux support and the lively discussions about anything and everything. Moreover, I would like to thank Corinne Eymann-Baier, Silja Kiriyanthan, Ketut Fundana, Tahir Majeed, Simon Pezold, Adrian Schneider, Francesco Santini, Zarko Celicanin, Stephan Wyder, Silvio Pflugi and Mathias Griessen for their participation in discussions on a range of scientific topics and much more.

Last and therefore most important, I would like to close by thanking my family. I thank my parents for having supported me throughout my studies. Without their support and confidence, I would not have been where I am today.

Summary

High Intensity Focused Ultrasound is an emerging non-invasive technology for the precise thermal ablation of pathological tissue deep within the body. The fitful, respiratory-induced motion of abdominal organs, such as of the liver, renders targeting challenging. The work in hand describes methods for imaging, modelling and managing respiratory-induced organ motion. The main objective is to enable 3D motion prediction of liver tumours for the treatment with Magnetic Resonance guided High Intensity Focused Ultrasound (MRgHIFU).

To model and predict respiratory motion, the liver motion is initially observed in 3D space. Fast acquired 2D magnetic resonance images are retrospectively reconstructed to time-resolved volumes, thus called 4DMRI (3D + time). From these volumes, dense deformation fields describing the motion from time-step to time-step are extracted using an intensity-based non-rigid registration algorithm. 4DMRI sequences of 20 subjects, providing long-term recordings of the variability in liver motion under free breathing, serve as the basis for this study.

Based on the obtained motion data, three main types of models were investigated and evaluated in clinically relevant scenarios. In particular, subject-specific motion models, inter-subject population-based motion models and the combination of both are compared in comprehensive studies. The analysis of the prediction experiments showed that statistical models based on Principal Component Analysis are well suited to describe the motion of a single subject as well as of a population of different and unobserved subjects. In order to enable target prediction, the respiratory state of the respective organ was tracked in near-real-time and a temporal prediction of its future position is estimated. The time span provided by the prediction is used to calculate the new target position and to readjust the treatment focus. In addition, novel methods for faster acquisition of subject-specific 3D data based on a manifold learner are presented and compared to the state-of-the art 4DMRI method.

The developed methods provide motion compensation techniques for the non-invasive and radiation-free treatment of pathological tissue in moving abdominal organs for MRgHIFU.

Zusammenfassung

High Intensity Focused Ultrasound ist eine aufkommende, nicht-invasive Technologie für die präzise thermische Zerstörung von pathologischem Gewebe im Körper. Die unregelmäßige ateminduzierte Bewegung der Unterleibsorgane, wie z.B. im Fall der Leber, macht genaues Zielen anspruchsvoll. Die vorliegende Arbeit beschreibt Verfahren zur Bildgebung, Modellierung und zur Regelung ateminduzierter Organbewegung. Das Hauptziel besteht darin, 3D Zielvorhersagen für die Behandlung von Lebertumoren mittels Magnetic Resonance guided High Intensity Focused Ultrasound (MRgHIFU) zu ermöglichen.

Um die Atembewegung modellieren und vorhersagen zu können, wird die Bewegung der Leber zuerst im dreidimensionalen Raum beobachtet. Schnell aufgenommene 2D-Magnetresonanz-Bilder wurden dabei rückwirkend zu Volumen mit sowohl guter zeitlicher als auch räumlicher Auflösung, daher 4DMRI (3D + Zeit) genannt, rekonstruiert. Aus diesen Volumen werden Deformationsfelder, welche die Bewegung von Zeitschritt zu Zeitschritt beschreiben, mit einem intensitätsbasierten, nicht-starren Registrierungsalgorithmus extrahiert. 4DMRI-Sequenzen von 20 Probanden, welche Langzeitaufzeichnungen von der Variabilität der Leberbewegung beinhalten, dienen als Grundlage für diese Studie.

Basierend auf den gewonnenen Bewegungsdaten wurden drei Arten von Modellen in klinisch relevanten Szenarien untersucht und evaluiert. Personen-spezifische Bewegungsmodelle, populationsbasierende Bewegungsmodelle und die Kombination beider wurden in umfassenden Studien verglichen. Die Analyse der Vorhersage-Experimente zeigte, dass statistische Modelle basierend auf Hauptkomponentenanalyse gut geeignet sind, um die Bewegung einer einzelnen Person sowie einer Population von unterschiedlichen und unbeobachteten Personen zu beschreiben. Die Bewegungsvorhersage basiert auf der Abschätzung der Organposition, welche fast in Echtzeit verfolgt wird. Die durch die Vorhersage bereitgestellte Zeitspanne wird verwendet, um die neue Zielposition zu berechnen und den Behandlungsfokus auszurichten. Darüber hinaus werden neue Methoden zur schnelleren Erfassung patienten-spezifischer 3D-Daten und deren Rekonstruktion vorgestellt und mit der gängigen 4DMRI-Methode verglichen.

Die entwickelten Methoden beschreiben Techniken zur nichtinvasiven und strahlungsfreien Behandlung von krankhaftem Gewebe in bewegten Unterleibsorganen mittels MRgHIFU.

Contents

Preface	i
Summary	iii
Zusammenfassung	v
1 Introduction	1
1.1 Motivation	2
1.2 Problem Statement	4
1.3 Research Methodology	5
1.4 Structure of the Thesis	6
2 Background	7
2.1 Imaging Respiratory-Induced Organ Motion	8
2.2 4DMRI in Clinical Applications	9
2.3 3D Volume Reconstruction	12
2.3.1 Stacking based on Template Matching	13
2.3.2 Stacking based on PCA	13
2.4 Navigator-less 3D Volume Reconstruction	14
2.5 Volume Registration	14
2.6 Extraction of the Breathing Signal	15
2.7 Motion Models	16
2.8 Database	17
2.9 MRgHIFU Setup	20
3 3D Organ Motion Prediction for MRgHIFU	25
4 Model-Based Respiratory Motion Compensation in MRgHIFU	35
5 Fast Manifold-based 4DMRI for Patient-Specific Organ Motion Modelling in MRgHIFU	47

6 Discussion and Conclusion	65
6.1 Answers to the Research Questions	66
6.2 Future Research	68
Curriculum Vitae	69
List of Publications	71

1 Introduction

Contents

Nowadays, minimally-invasive and non-invasive treatment methods are well established and used in many surgical application. The benefits of such treatments are less pain, faster recovery and, therefore, less operative trauma for the patient than equivalent invasive procedures. The application of non-invasive techniques may result in shorter hospital stays or allows outpatient treatments. Magnetic Resonance guided High Intensity Focused Ultrasound, for example, has the unique capability to deposit sharply localised energy deep within the body, producing thermal ablation without damaging the surrounding tissue. Precise targeting, however, demands for exact knowledge of the target position. When treating abdominal organs such as the liver, the fitful respiratory-induced organ motion renders targeting challenging. This thesis proposes different approaches that allow the compensation of respiratory-induced organ motion. This chapter introduces the research field as well as the problem statement of the thesis.

Outline

Section 1.1 presents an overview of the research topic, the state-of-the-art methods and the need for research. This leads to the actual problem statement that is formulated in two research questions as presented in Section 1.2. Section 1.3 outlines the research methodology that is employed to investigate and answer the stated research questions. The chapter concludes by an overview of the structure of the thesis.

1.1 Motivation

Cancer is a leading cause of death worldwide, accounting for 7.6 million deaths (around 13% of all deaths) in 2008 [1], whereby liver cancer is the third most common cancer. Liver cancers are malignant tumours growing inside or on the surface of the liver. The correct and timely treatment can mean the difference between life and death. Available treatment methods include surgery, chemotherapy, radiation therapy and hyperthermia. Currently, research is heading towards non-invasive treatment methods, such as Magnetic Resonance guided High Intensity Focused Ultrasound (MRgHIFU). Focused ultrasound has the unique capability to deposit sharply localised energy deep within the body producing thermal ablation without damaging the surrounding tissue as illustrated in Figure 1.2. The increasing capabilities of more accurate targeting demands for more exact knowledge of the target position. However, the fitful respiratory-induced motion of the liver (see Figure 1.1) renders non-invasive treatments challenging.

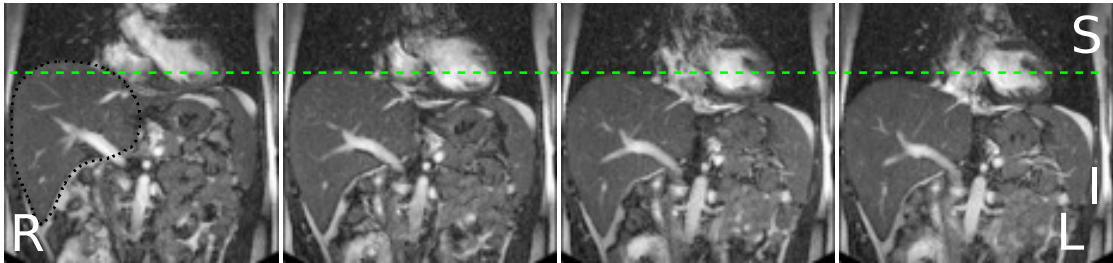


Figure 1.1: Coronal view of the liver (*dotted line*) under respiratory-induced motion moving from exhalation state (*left*) to inhalation state (*right*).

The work in hand is part of a research project aiming at developing motion compensation techniques for non-invasive and radiation-free ablation of localised cancers in moving organs using an MRgHIFU device. From a clinical perspective, MRgHIFU has several advantages compared to traditional surgery. Since no ionising radiation is applied, the therapy can be repeated as often as necessary. As the intervention is non-invasive pain is minimal and recovery is faster. MRgHIFU even has the potential for an out-patient treatment. Besides, the procedure cost is low as compared to traditional surgical methods. Focusing is achieved by constructive interference of sound waves emitted from *e.g.* a 256 element phased array transducer operating at 1 MHz. High Intensity Focused Ultrasound (HIFU) is a hyperthermia therapy that uses temperature to treat diseases. At the focal point, the temperature can rise to levels of 65°C to 85°C, destroying the diseased tissue by coagulation necrosis. Phased-array HIFU devices of the latest generation reach oblate ellipsoid lesion sizes of around 2 mm width and 4 mm length depending on exposure time, acoustic power and sonication pattern [2]. Beam spot steering ranges up to 70 mm [3] and allows the continuous application of

HIFU energy to the mobile target without interruption, which is essential to achieve the targeted thermal dose required for complete coagulation within typical exposure times of approximately 20 seconds. Another approach to handle breathing-induced motion is respiratory gating, with the advantage that neither real-time tracking of the exact target position nor real-time beam steering capabilities of the HIFU system is required. But respiratory gating significantly reduces the overall duty cycle for the sonication process, merely leaving a temporal window of 1-2 s within the respiratory cycle. Moreover, the high perfusion rate of the liver can lead to a strong heat evacuation which limits a sufficient temperature increase to achieve necrosis in large target volumes [4,5]. Performing motion compensation, one has to expect target movements with a main component in superior-inferior (S - I) direction, usually in the range of 5-25 mm for quiet breathing [6,7]. Additional motion in anterior-posterior (A - P) and left-right (L - R) direction in the range of 1-12 mm and 1-3 mm, respectively, as well as non-rigid deformations of up to 20 mm as quantified by [7] has to be compensated. Organ deformation, probably caused by intestinal activity and muscle relaxation occurring over longer time scales, the so called drift, also have to be taken into account [8].

Besides the flexible steering capabilities, a key benefit of HIFU over many other forms of focused energy therapy, such as radiation or radio surgery, is, that the passage of ultrasound energy through intervening tissue has no apparent cumulative effect on that tissue. To this end, the treatment can be repeated without causing collateral damage to neighbouring healthy tissue. Magnetic Resonance (MR) Imaging, on the other hand, offers excellent anatomical tissue contrast allowing to observe, study and plan the targeting of the tumour, thus, allowing for spatial guidance of the therapy. In addition, MR thermometry provides outstanding capabilities for on-line temperature control during sonication [9]. This is currently the only available technique for non-invasive temperature measurement within biological tissue. Temperature maps can be used as an active feedback control of the HIFU beam in order to quantify the thermal dose delivered to the target with a typical resolution of 1 mm (spatial), 1 second (temporal) and 1°C in temperature, respectively.

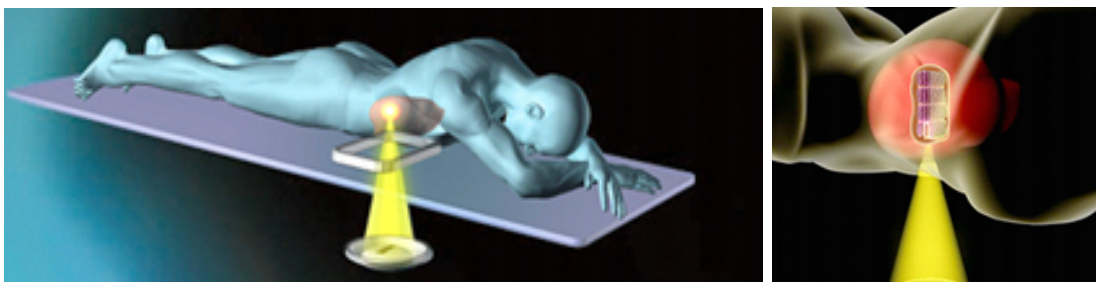


Figure 1.2: Principle of HIFU ablation in the liver (Courtesy of www.biomaxx-sys.com).

The drawback of using MR thermometry in MRgHIFU is, though, that the available scan time is mainly required for the temperature feedback control. Although the patient is located within the MR system during sonication, there's only very little scan time left to track the target. Therefore, a key issue of the procedure is, that extensive 3D target tracking is not feasible during treatment. Several approaches have been proposed in the literature to manage respiratory-induced organ motion. The state-of-the-art methods, such as proposed by [10], gather a 2D motion atlas in an initial learning phase, which is used during sonication to correct the target position. The main drawback of the method is, however, that it neglects the out-of-plane motion that can amount to a couple of millimetres in L-R direction. In [11], a method was proposed whereby an MR navigator echo enhanced by a Kalman filter is used to robustly track the moving organ. Using only a one-dimensional navigator, only the rigid translational motion is determined. Again, both cyclic non-rigid deformation and drift are not detected and compensated.

A real-time tracking method that observes the target on a 2D image plane combined with a perpendicularly acquired pencil beam navigator, finally obtaining 3D information of the targets trajectories, was proposed in [12]. Since no 3D ground truth data was available, the performance of the method was evaluated by comparing the temperature distributions obtained after 60 s of HIFU application with and without motion compensation, resulting in higher final temperatures in the target area with enabled motion compensation. This method was tested in vivo with kidneys of ventilated pigs, both following a regular and stable breathing pattern. A review of recent technological advances in MR-temperature mapping of moving organs, in motion compensation of the HIFU beam, and in volumetric ablation and feedback control strategies, is given by [4]. It is noteworthy that none of the state-of-the-art approaches have the capabilities to examine their proposed approach on real 4D data acquired over long time-scales.

1.2 Problem Statement

Non-MR-based external surrogate markers (*e.g.* optical tracker, breathing belt) or fast MR-based 1D navigator pencil beams [12, 13] would be ideal for real-time tracking of the respiratory-induced organ motion, demanding for none or only a marginal MR scan-time. In addition, tracking surrogate markers provides low-lag breathing signals to predict the actual position of the tumour. Since the motion of the target occurs in 3D space, a one-dimensional tracked signal might not be sufficient for the accurate prediction of the target location. Due to the limited possibilities of target tracking caused by the restricted MR imaging time, the prediction of 3D organ motion has to be based on the measurement of a 1D signal. The connection between the 1D signal and the 3D motion of the organ must be given by underlying models. Since there's no scan time available during sonication, the model must be acquired prior to the actual treatment. Such a model can also be based on population statistics, where no subject-specific mo-

tion data is required. The creation of various models is the specific aim of this thesis. From the problem statement above, we can derive a general research question:

Research question 1: *How accurate can the 3D position of a target be predicted, given that the respiratory state of the liver can be tracked in 1D?*

The objective to solve is twofold: First, in order to compensate for the system lag, that is composed of the time needed to track the position of the surrogate and for data processing to estimate the new target position, one has to estimate the future course of the measured signal. In our application, the 1D signal describes the respiratory state of the liver that is measured by tracking the height of the diaphragm. Note, that this 1D signal is only describing the motion of one tracked point located at the diaphragm. Therefore, having an estimate of the future respiratory state, the second task is then to associate the motion of the tracked point with the motion of the entire liver. All the motion modelling approaches described in this thesis builds on the work of Martin v. Siebenthal [14], who developed the 4DMRI method in the scope of his dissertation. 4DMRI was especially designed to acquire 4D data over tens of minutes, with the specific aim to study the long-term intra-subject variability of organ motion of free breathing awake patients, which is still poorly understood. As the 4DMRI method was developed for long-term studies, it is not designed for a clinical environment, where short acquisition times and fast data processing are crucial. This brings us to a second important question:

Research question 2: *How much pre-operative data is required to construct models suitable for motion prediction?*

To answer these two question, we followed the research methodology presented in the next section.

1.3 Research Methodology

The research methodology followed in this thesis consists of imaging and observing the organ motion in an initial learning phase, constructing motion models, employing the models for prediction and finally, evaluating the prediction based on the available ground truth motion. An overview is outlined in Figure 1.3. It consists of six major parts, highlighted by the boxes. The respiratory-induced organ motion is first imaged by the 4DMRI acquisition sequence similar to the one proposed in [8]. The acquired 2D slices are then retrospectively reconstructed or stacked to 3D volumes with both a high temporal and spatial resolution. The obtained stacks contain a variety of respiratory states, observed during the 4DMRI acquisition. By means of 3D non-rigid registration [15], vector fields, describing the organ motion from time step to time step are extracted. These vector fields serve as the basic data to construct the motion models as well as the

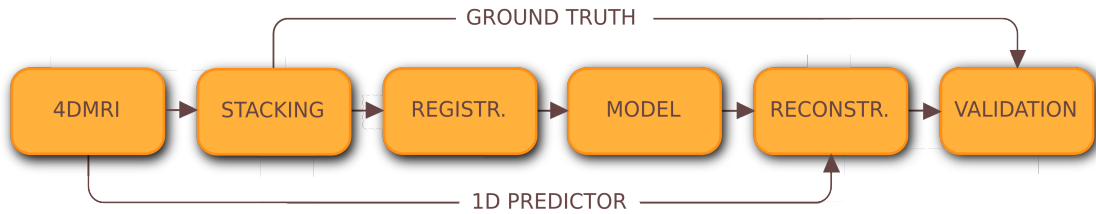


Figure 1.3: Schematic flowchart of the overall motion prediction procedure. Based on 4DMRI data, volumes are retrospectively reconstructed, registered, used to construct motion models and finally, evaluated on ground truth data. The breathing signal is extracted from the 4DMRI sequence and used for the motion reconstruction.

ground truth to evaluate the prediction performance of the models. The breathing signal is extracted from the 4DMRI data as well and used for the temporal prediction.

This work suggests possible ways to manage the variability of respiratory-induced organ motion aiming at a safe and efficient treatment in MRgHIFU. The approaches proposed in this thesis were developed more and more towards the direction of clinically applicable treatment techniques.

1.4 Structure of the Thesis

The remainder of the thesis consists of two main parts. The first part introduces the research area and delivers the required background information for a clear understanding of the following chapters. The second part consists of the conducted studies integrated as separate chapters, comprising the elaborated publications. Below, the contents of each chapter is briefly discussed. This chapter presented the motivation, problem statement, research question, research methodology and the structure of the thesis. In Chapter 2, background information of the employed methods is delivered. Chapters 3, 4, 5 attempt to answer the research questions. The prediction performance of subject-specific motion models in combination with population-based drift models (in Chapter 3) and pure population-based motion models (in Chapter 4) is investigated. In Chapter 5, our focus shifts towards fast acquisition of pure subject-specific motion models with the specific aim of being applicable in a clinical environment. Chapter 6 concludes the thesis and provides directions for future research.

2 Background

- Contents** This chapter provides background information about the preparatory work, the applied techniques and the database used for the experiments. It should help the reader to understand the remainder of the thesis. 4DMRI was developed for the long-term study of respiratory-induced organ motion and is therefore well suited to investigate *research question 1* as stated in Chapter 2. The fusion of 4DMRI with MRgHIFU to allow for a clinical applicable treatment demands for shorter 4DMRI acquisition times and fewer manual user-inputs as compared to the state-of-the-art 4DMRI [8]. In general, the clinical applicability is the central theme of this thesis. From the initial question of how accurate the liver's motion can be predicted, the focus of the study moves towards the question of how much data is actually needed for accurate prediction.
- Outline** Sections 2.1 and 2.2 introduce 4DMRI and its limitations for the clinical use. Section 2.3 presents alternative approaches to retrospectively reconstruct 3D volumes based on 4DMRI data and Section 2.5 introduces the applied registration algorithm used to extract the deformation fields. Section 2.6 describes the extraction of the breathing signal from the 4DMRI sequence. The obtained deformation fields provide the basis for both the motion modelling, as is described in Section 2.7, and the evaluation of the experiments. Finally, Section 2.8 presents the database used for the experiments and Section 2.9 presents a prototype of an MRgHIFU system.

2.1 Imaging Respiratory-Induced Organ Motion

Although MR techniques are constantly evolving, it is not possible to capture large 3D volumes of high resolution image data in real-time without suffering from a poor signal-to-noise ratio (SNR). To obtain 3D volumes of, for example, a stationary organ, 2D slices are acquired consecutively at different slice positions covering the volume of interest (1 pass = 1 *dynamic*) and are stacked to a smooth 3D volume. But when imaging moving abdominal organs such as the liver, non-smooth reconstructions would be obtained by applying simple slice stacking. Due to the respiratory-induced volume changes of the lung, the abdominal organs and the lung vary in shape and position. These changes have to be compensated for in order to obtain proper volumes. One reliable possibility to obtain smooth 3D volume reconstructions over time, thus 4D, is to retrospectively sort images captured during multiple *dynamics* applying internal gating [16]. This method called 4DMRI allows to capture and observe the motion of moving abdominal organs during free breathing for hours. Hereby, a fast interleaved imaging sequence consisting of *data slices* and *navigator slices* is used (see Chapter 5, Figure 2(a), 2(b)). While the *data slices* are acquired at different slice positions to cover the volume of interest, the *navigator slices* are acquired at a fixed position. Each *data slice* is thus embraced by two *navigator slices* which allows to robustly identify the respiratory state of the embraced *data slice*. Based on a similarity measure between the embracing *navigator slices*, complete 3D volumes are retrospectively reconstructed. For more detailed information regarding 4DMRI imaging, we refer the reader to the work of von Siebenthal *et al.* [8, 14]. Although the method does not capture the full organ at once, accurate 3D volumes under free breathing, comparable to breath-hold images, are achieved. For the first time, insight into the variability of respiratory-induced organ motion of the liver could be provided. The possibility to observe 3D organ motion over long time scales opened up the field for a wide range of studies and applications such as those proposed by [8, 14, 17–20].

Since the goal of this dissertation is to compensate for the liver motion of free-breathing patients, we first take a look at the variability of the respiratory-induced organ motion and at the characteristics of the underlying breathing pattern. Depending on the patient, *e.g.* his health situation and anatomy, the respiratory pattern and corresponding organ motion, respectively, can vary drastically in amplitude, cycle duration and regularity. Figure 2.1(a) shows an example of a regular breathing pattern with a stable amplitude of about 14 mm and a cycle duration of about 5 s. In contrast, Figure 2.1(b) shows an irregular breathing pattern with changes in amplitude, cycle duration and exhalation position, ranging from 4-9 mm, 3-5 s and 3 mm, respectively. Fitful breathing not only complicates the prediction of the motion, but also the quality of the reconstructed volumes. 4DMRI follows a retrospective slice stacking approach, which implicates that every slice position has to be captured at every respiratory state in order to reconstruct

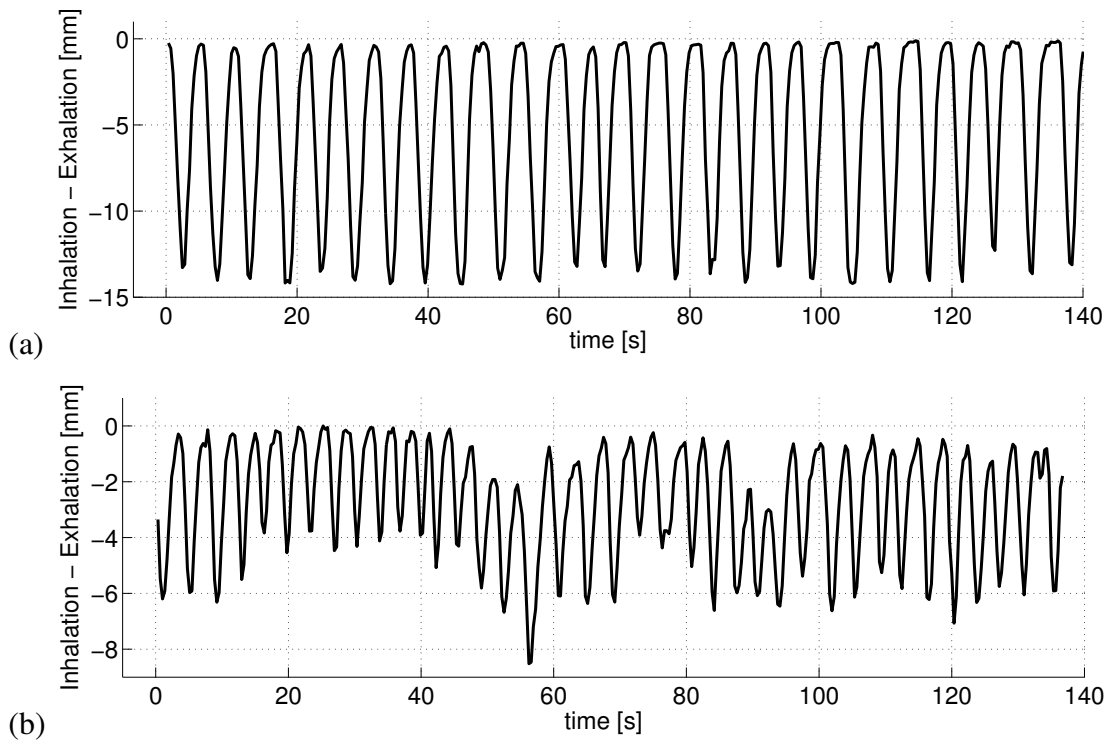


Figure 2.1: Variability of superior-inferior diaphragm motion of two subjects showing large differences in the breathing pattern. (a) Regular breathing pattern with an amplitude of approximately 14 mm and breathing period of about 5 s. (b) Irregular breathing pattern with an amplitude of up to 9 mm, a breathing period of about 3 s and, in addition, a varying exhalation position.

a complete volume at the respective breathing depth. Let us take, for example, the case where a volume consisting of 25 slices in the mid-inhalation phase should be imaged. This implies that in the best case, at least 25 respiratory cycles have to be acquired in order to reconstruct one complete stack at the respective breathing state. The more irregular a breathing pattern is, the more unlikely it is that complete volumes for all occurring breathing depths can be reconstructed. In any case, a certain amount of data or 4DMRI acquisition time is required to ensure suitable volume reconstructions.

2.2 4DMRI in Clinical Applications

To achieve high quality stacks, such as presented by [8], an immense amount of data has to be acquired. In a 35 minute scan, the volume of interest is imaged around 200-

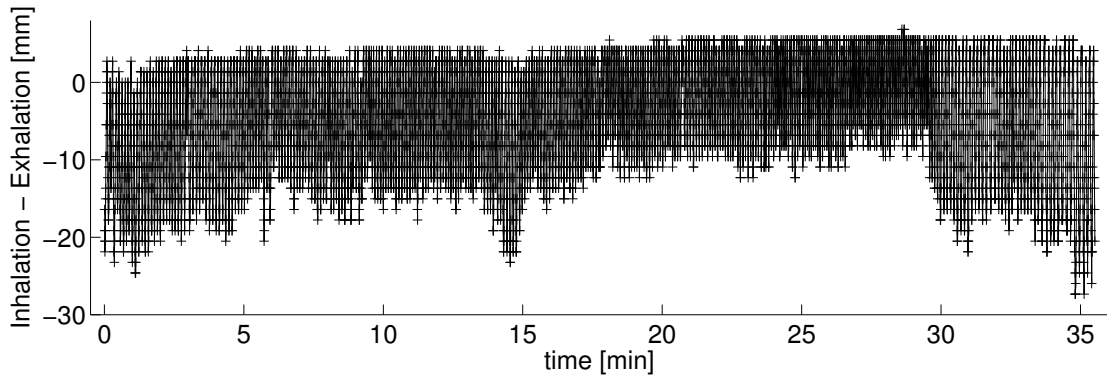


Figure 2.2: Respiratory breathing signal of Subject 1 extracted from 35 minutes of 4DMRI acquisition under free breathing. The breathing pattern shows a stable exhalation position with a small upwards drift and a varying inhalation position.

300 times (*dynamics*), whereby the *navigator slice* is acquired up to 6000 times. This procedure might be acceptable for studies of healthy volunteers to investigate the long-term behaviour of a moving organ, but not for the application in a clinical environment, treating patients with advanced stage cancer. For the clinical applicability of 4DMRI, shorter acquisition times of no longer than 10 minutes should be considered. With decreasing acquisition time, certain limitations of the 4DMRI approach become noticeable. Due to the irregularity of breathing, not the entire range of respiratory breathing depths are equivalently sampled by the 4DMRI scan. The frame rate of *data slices* is around 350 ms, while the duration of a respiratory cycle takes around 4 seconds, which corresponds to a sampling rate of approximately 11 images per respiratory cycle. Figure 2.2 shows a breathing pattern, extracted by template matching (see Section 2.6), from a 35 minute 4DMRI acquisition. While the exhalation baseline is typically very stable, the full inhalation depth is strongly varying. For deep inhalation states, occurring at the beginning of the example sequence, at around 15 minutes and 35 minutes, it is not possible to reconstruct a complete stack. This issue is best illustrated with an example. Figure 2.3(a) and Figure 2.3(b) show very typical distributions of the *navigator slices* and of the corresponding *data slices*. The figures reveal the counts per bin for the *navigator slices* and the bin occupancy for a given slice position and breathing state, respectively. The white fields indicate unoccupied bins. The distribution typically shows a global maximum at the exhalation state and a local maximum at the inhalation state positions. The distribution of the *data slices* in Figure 2.3(b) and Figure 2.3(d) shows that not all bins are occupied. Rare and very deep breathing cycles can therefore not be properly reconstructed. In such a case, the best matching slice which, for deep inhalation states, means *data slices* from a shallower breathing state, is taken for the reconstruction and are, hence, not representative for the deepest breathing state. As a

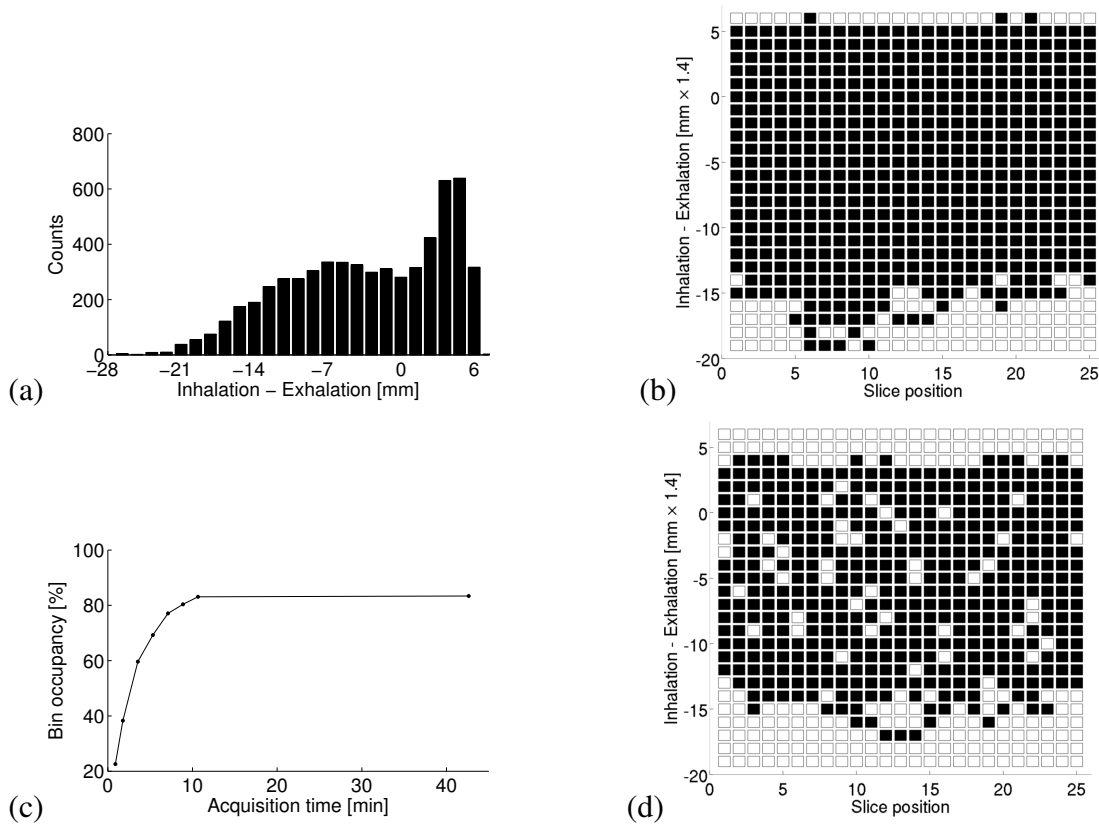


Figure 2.3: (a) Distribution of the *navigator slices* given the inhalation depth and (b) corresponding bin occupancy of *data slices* for given slice positions and inhalation depths after 35 minutes of 4DMRI acquisition of Subject 1. (c) Percentage of complete stacks shown for Subjects 1 given the respective acquisition times. (d) Bin occupancy after 7 minutes of 4DMRI data acquisition given for Subject 1.

first rule, one can observe that more data is available in the exhalation state than in the inhalation state (see Figure 2.3(a)), which results in a better stacking quality for exhalation stacks. To illustrate the effect of short acquisition times, Figure 2.3(d) shows the bin occupancy after a 7 minutes 4DMRI acquisition, where only a few respiratory states are completely filled. In such cases, the best matching *data slices* are taken from adjacent bins for the reconstruction, leading to less accurate reconstructions. Figure 2.3(c) shows the bin occupancy for different acquisition lengths calculated for Subject 1. It can be clearly seen that on average, after 10 minutes acquisition time, around 80 % of the observed breathing depths are fully captured. The occupancy is not further increasing with increasing acquisition time. This behaviour can be traced back to both the facts that new respiratory bins arise while the liver is drifting over time [8] as well as to the rare occurrence of very deep breathing cycles. As a second rule, one can state, that the

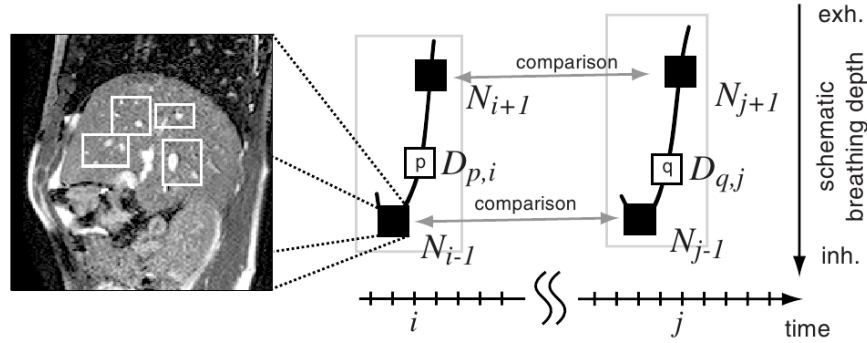


Figure 2.4: The correspondence of two *data slices* D acquired at different slice locations p, q and time points i, j is determined by comparing the embracing *navigator slices*. The example *navigator slice* shows four regions that were considered for image sorting (Courtesy of Martin von Siebenthal [8]).

longer the acquisition time, the higher the probability that all the observed breathing depths can be fully reconstructed. In general, one has to take into account that rarely occurring breathing depths are not reliably captured and reconstructed, even during long acquisition sessions.

2.3 3D Volume Reconstruction

The key idea of 4DMRI is the interleaved acquisition of *data slices* and *navigator slices*, whereof the *navigator slices* are dedicated to identify the respiratory state. *Data slices* of the same respiratory state are stacked together, resulting in one reconstructed volume per pair of *navigator slices*. Figure 2.4 illustrates the slice stacking approach proposed by [8]. Considering the *data slice* $D_{p,i}$ at position p and time point i , we want to find a *data slice* $D_{q,j}$ at an other sagittal position q and time point j , but at the same respiratory state. Therefore, we compare the embracing *navigator slices* N_{i-1} and N_{i+1} of $D_{p,i}$ to all *navigator slices* N_{j-1} and N_{j+1} of all possible candidate slices $D_{q,j}$. The comparison is based on a frame similarity measure of the embracing *navigator slices*. The frame similarity is determined by the cost function $C(i, j)$, which is small if the *data slices* $D_{p,i}$ and $D_{q,j}$ show the same respiratory state. Given a *data slice* $D_{q,i}$, the best matching *data slice* D_{q,j^*} is thus found by

$$j^* = \arg \min_j C(i, j). \quad (2.1)$$

We will shortly introduce the slice matching criteria and the resulting cost functions that we used for retrospective slice stacking.

2.3.1 Stacking based on Template Matching

Figure 2.4 illustrates a similarity measure based on template matching of four manually selected regions of interest, indicated by four white frames. Although template matching describes a rigid transformation, it is possible to detect the non-rigid motion among the templates. Due to the distribution of the templates (see Figure 2.4), organ deformation occurring during the respiratory cycle and drift occurring over longer time scales are well detected. As proposed by [8], we assume coordinate axes in anterior-posterior and inferior-superior direction. The coordinates x_i^k and y_i^k describe the centre of template k at time i , where K is the number of tracked regions. The cost function for the template matching approach is thus given by

$$C(i, j) = \sum_{k=1}^K \left\| \begin{pmatrix} x_{j-1}^k - x_{i-1}^k \\ y_{j-1}^k - y_{i-1}^k \end{pmatrix} + \begin{pmatrix} x_{j+1}^k - x_{i+1}^k \\ y_{j+1}^k - y_{i+1}^k \end{pmatrix} \right\| = \sum_{k=1}^K \left\| \Delta \vec{X}_{-1}^k + \Delta \vec{X}_{+1}^k \right\|. \quad (2.2)$$

The K displacements of the preceding \vec{X}_{-1}^k and of the subsequent \vec{X}_{+1}^k navigator frames are summed up by vector addition. Vector addition was chosen to enable the detection of *data slices* that are actually identical but were acquired during organ movements with different speeds. The drawback of vector addition is, though, that *data slices* acquired at a different breathing phase, *i.e.* ex- or inhalation, but with similar speed of movements are matched together. This can happen especially in the case, where only one region is tracked and used for comparison. While the method is easy to implement, a drawback is the required manual user input. In particular, it can be difficult to select templates showing vessels with high contrast that are visible during the entire respiratory cycle. Moreover, due to out of plane drifts, it can happen that vessels change in shape and size or even disappear during long acquisition sessions. To ensure a reliable selection of templates, the template tracking results have to be reviewed and the template selection has to be repeated in case of problems.

2.3.2 Stacking based on PCA

To avoid manual user input, Principle Component Analysis (PCA) can be applied to the *navigator slices* as is described in Chapter 5. After projecting all *navigator slices* into the PCA-space, every *navigator slice* can be described by its PCA coefficient vector \mathbf{c} . The cost function for PCA-stacking is, therefore, defined by the distance between the respective pairs of *navigator slices*:

$$C(i, j) = \|(\mathbf{c}_{i-1} - \mathbf{c}_{j-1}) + (\mathbf{c}_{i+1} - \mathbf{c}_{j+1})\| = \|\Delta \mathbf{c}_{-1} + \Delta \mathbf{c}_{+1}\|. \quad (2.3)$$

The coefficient vectors \mathbf{c}_{i-1} and \mathbf{c}_{i+1} describe the respective *navigator slices* N_{i-1} and N_{i+1} of *data slice* D_i in the navigator space.

2.4 Navigator-less 3D Volume Reconstruction

The current state-of-the-art 4DMRI method as described above, is based on the interleaved acquisition of *data slices* and corresponding *navigator slices*, which are needed to identify the respiratory state of the organ for the 3D reconstructions. This approach is very well suited for the detailed study of long-term organ deformations or drifts. When short acquisition times are desirable, as is for example the case in clinical applications, the present 4DMRI sequence can result in an undersampling of *data slices* as discussed in Section 2.2. One possibility to improve the present sequence is to avoid the acquisition of the *navigator slices*. Currently, 50% of all acquired images are only used for volume stacking, but are of no further interest. The possibility of retrospectively reconstructing volumes without the need of interleaved *navigator slices* does not only allow to half the acquisition time, but also double the amount of acquired data within a given time period and, hence, double the acquisition speed and frame rate. *Navigator-less* stacking is described and evaluated in detail in Chapter 5.

2.5 Volume Registration

After the reconstruction of the stacks, we extract the deformation fields that describe the motion from one stack to the subsequent stack. This is done by the B-spline based non-rigid registration algorithm proposed by [15]. The current implementation of the applied registration algorithm is, however, not able to handle discontinuous motion. As it occurs at the organ boundaries, where for example the liver slides along the chest wall. Since we are only interested in the motion inside the liver, a template liver was manually segmented and registered upon all other volume reconstructions. During registration, the masked voxels outside the liver were ignored by the algorithm. This results in deformation fields which are only valid within the liver. Figure 2.5 shows the result of registering the manually segmented liver template to a complete respiratory cycle. The resulting deformation fields (*white*) are overlaid with the reconstructed volume, showing the liver's motion in all three spatial directions. The regions covered with deformation fields indicate the outline of the manually segmented template liver. The deformation fields are either extracted at an isotropic grid with 15 mm resolution (Chapter 3 and Chapter 5) or at an inter-subject corresponding grid (Chapter 4). After the extraction of the deformation fields, the motion of the organ is described by deformations applied to the grid points. As discontinuous displacement fields are quite common in the medical field, discontinuity-preserving registration is an important research topic. Novel methods, such as [21], integrate registration and motion segmentation in a variational framework to handle discontinuous motion fields.

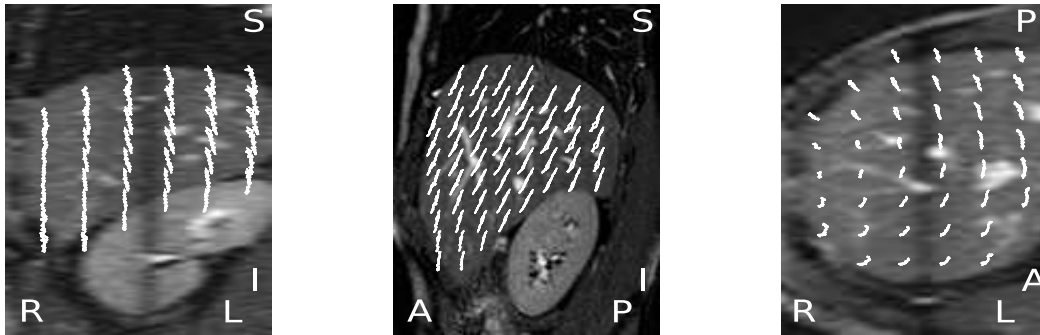


Figure 2.5: Exemplary deformation fields extracted using B-spline based non-rigid registration, representing the liver’s motion during one respiratory cycle.

2.6 Extraction of the Breathing Signal

In order to compensate respiratory-induced organ motion, either the current target position or surrogate markers resolving the target position have to be tracked. Based on these measurements, an estimate of the future position has to be formed and the treatment focus adapted to the newly calculated target position. The time lag between target tracking and repositioning of the treatment focus has to be bridged by a signal prediction algorithm. Taking a step back and having a look at the flowchart (Figure 1.3) presented at the beginning in Section 1.3, we see an arrow labelled as “1D PREDICTOR”. This arrow indicates that the 1D signal prediction is based on information gathered from the 4DMRI sequence, processed and later used for the evaluation of the motion reconstruction process. As no surrogate marker was deployed and, therefore, no breathing signal was acquired simultaneously during the acquisition sessions, the breathing signal is extracted from the 4DMRI sequence. This is done by tracking a manually selected region placed at the diaphragm of a *navigator slice* throughout all subsequent *navigator slices*, as shown in Figure 2.6(a). We used template matching with cross-correlation as the similarity criterion for tracking. As the volumes are reconstructed at the time points between the *navigator slices*, the extracted breathing signal is linearly interpolated to obtain the respective respiratory state at the time points of the reconstructions. In all of the performed experiments, predictions are based on this breathing signal. The respective breath prediction techniques are described in detail in the Chapters 3, 4 and 5. Having an estimate of the future position of the template, the template centre (+) is assigned to the closest point (●) of the point grid (○), which is used for the registration, as depicted in Figure 2.6(b). The closest grid point is found by measuring the Euclidean distance.

So far, an estimate of the future displacement of the tracked template and, therefore, of the adopted grid point (●) is known. To predict the motion of the entire liver, the rigid

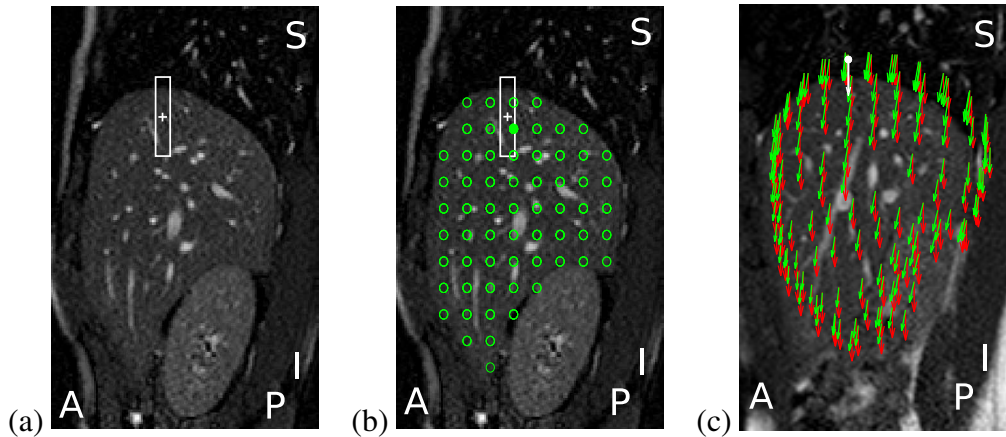


Figure 2.6: (a) Extraction of the breathing signal from *navigator slices* with a manually selected region at the diaphragm used for template matching. (b) Selection of closest model grid point (●) with respect to the centre of the selected template (+). (c) Motion of tracked grid point (*white arrow*), corresponding rigid correction (↓) and non-rigid motion correction provided by the motion model (↓).

shift (↓) could be assigned to all other grid points. This results in the simple case of the rigid correction (↓) visualised in Figure 2.6(c). Applying the motion models presented in Chapter 4 and Chapter 5, more exact motion prediction can be achieved. Given the rigid shift of the tracked template, a non-rigid motion field (↓) is reconstructed and used for the prediction.

2.7 Motion Models

Different types of motion models are investigated and presented in the Chapters 3, 4 and 5. In particular, three types of motion models are further investigated. In Chapter 3, a patient-specific motion model combined with a population-based statistical drift model based on PCA is presented. The results of this method are very convincing. Especially the combination of patient-specific breathing motion data with population-based drift data makes the method suitable for long-term predictions, as organ drifts are incorporated. This method, however, involves long acquisition- and data processing times to create the motion model, which renders the method unsuitable for clinical use.

To overcome this issue, a motion compensation approach based on population statistics is proposed in Chapter 4. Using a population-based statistical model has the advantage that only little patient-specific data, namely a breath-hold scan, has to be acquired. Aligning the new patient liver and the population-based statistical model is the major difficulty of the method, though. An alignment is only possible if the treatment of the

patient occurs in the same pose as for which the model was constructed. This brings along that even an additional device, such as an ultrasound probe that is pressed into the abdomen to track the breathing state, can invalidate the model. Such a model is acquired for a specific organ and is therefore not applicable to other organs. Even if a motion model for each individual type of organ, such as the lung, kidney and liver, is available, the question, whether the models cover the motion of the entire population, remains open. Nevertheless, it is an elegant way to handle respiratory-induced organ motion.

Based on the issues raised above, a fast and general applicable model to manage respiratory-induced organ motion would thus be ideal for clinical use. A subject-specific modelling approach is the only way to consider all the possible treatment scenarios and individual characteristics of each patient. Therefore, the key is to optimise the subject-specific approach in terms of speed without sacrificing precision. Chapter 5 proposes such fast acquisition methods and modelling approaches. The chapters follow the chronological order of publications.

2.8 Database

Figure 2.7 shows the exhalation surfaces of 20 healthy subjects. 4DMRI sequences of these 20 subjects form the basis of all experiments performed in this thesis. These sequences were acquired on a 1.5 T Philips Achieva whole body MR system in the scope of the dissertation of Martin von Siebenthal. We refer the reader to the respective thesis [14] for detailed information of the acquisition sequence. For each of the 20 subjects, the liver was manually segmented. The livers are numbered from 1 to 20, whereby the datasets of Subjects 1, 2, 3, 4, 5, 6, 7, 8, 11, 12, 13 and 14 are used in Chapter 3. In Chapters 4 and 5, the database was extended to all 20 subjects. Figure 2.8 gives an overview of the average motion we have to deal with within this thesis. The greyscale values of the liver surface indicate the mean motion shown for each spatial direction. In the datasets, we observe mean motion of 1.4 mm in A-P direction, 11.7 mm in I-S direction and 0.3 mm in L-R direction. The mean motion was obtained by averaging the respective motion components over all available time steps for the given grid points. The shape of the liver in Figure 2.8 is the average shape of the 20 liver shapes shown in Figure 2.7. The overall motion is composed of respiratory-induced motion and additional drifts. Figure 2.9 shows the average motion of the exhalation position, thus the drift, and helps to understand which kind of motion contributes to the overall motion. Comparing Figure 2.8(b) and Figure 2.9(b), it is clearly visible that the large motion at the tip of the liver is caused by the drift.

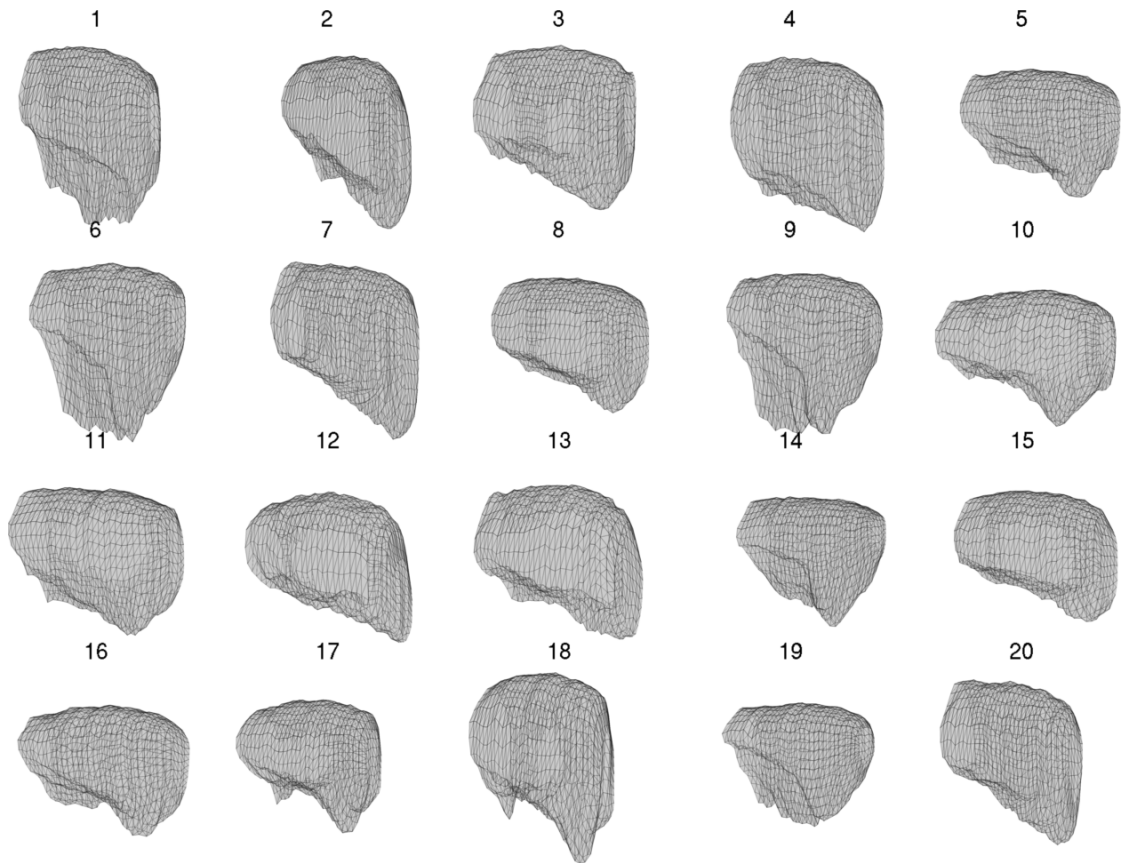


Figure 2.7: Surfaces of all 20 liver subjects after manual segmentation. The datasets of Subjects 1, 2, 3, 4, 5, 6, 7, 8, 11, 12, 13 and 14 are used in Chapter 3.

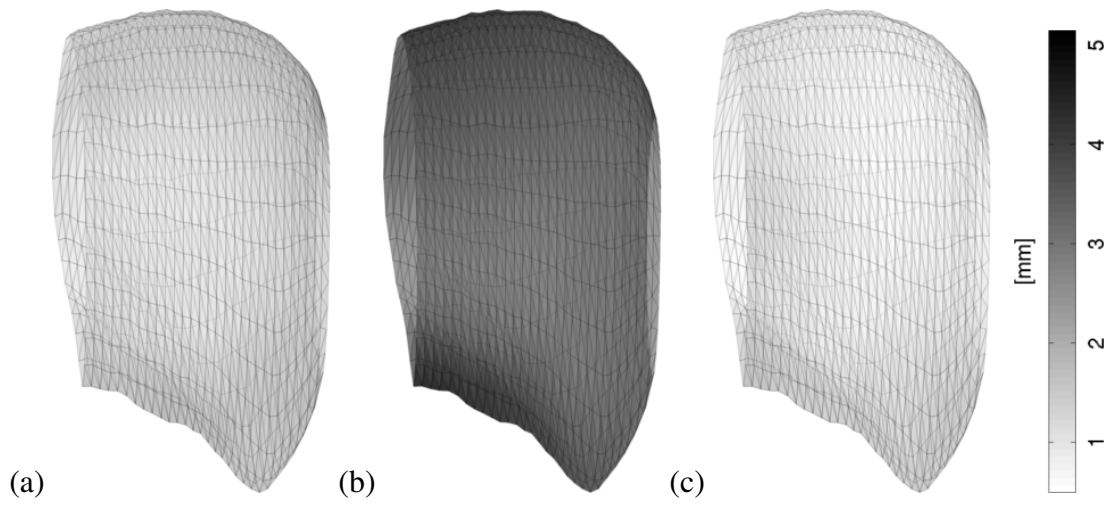


Figure 2.8: Liver surfaces showing the mean motion for (a) A - P, (b) I - S and (c) L - R direction with an average motion amplitude of 1.4 mm, 3.9 mm and 1 mm, respectively.

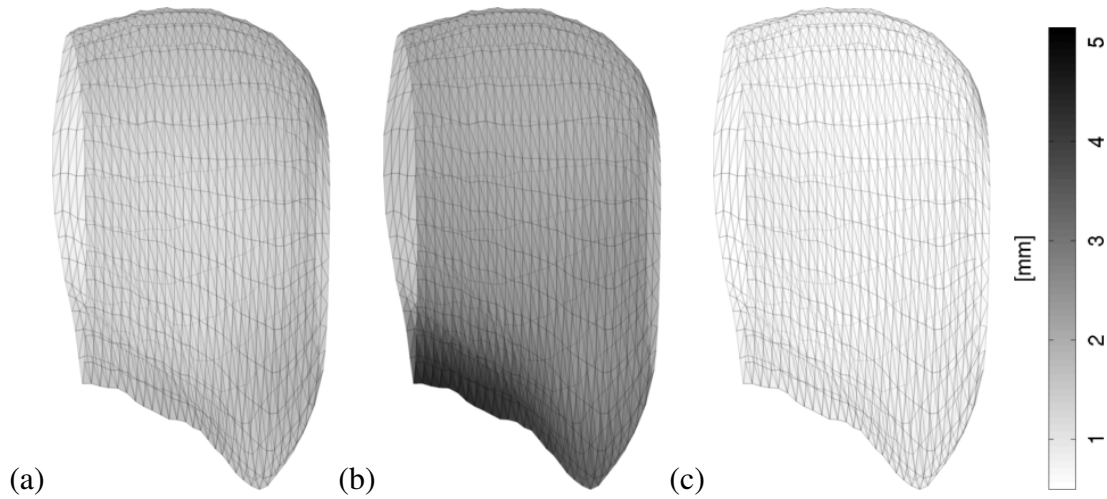


Figure 2.9: Liver surfaces showing the mean drift shown for (a) A - P, (b) I - S and (c) L - R direction with an average drift amplitude of 1.4 mm, 2.6 mm and 0.8 mm, respectively.

2.9 MRgHIFU Setup

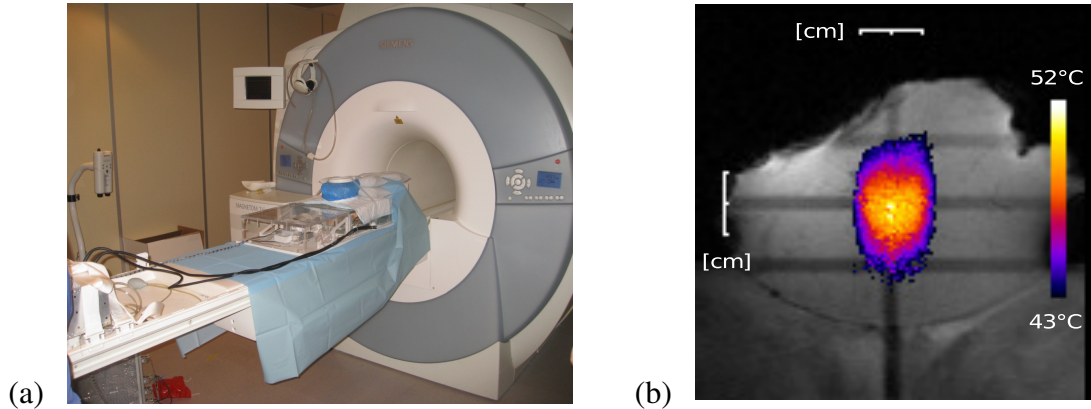


Figure 2.10: (a) Experimental setup: MR scanning system and MR-compatible HIFU prototype at the University Hospital Geneva (HUG). (b) Example of fast HIFU ablation ($5 \text{ cm}^3/\text{min}$) of 120 seconds duration in a non-moving target. The image shows a slice through the sonicated volume of size $2 \text{ cm} \times 2.5 \text{ cm} \times 2 \text{ cm} \approx 10 \text{ cm}^3$.

In the following three chapters, motion compensation techniques to manage respiratory-induced organ motion are proposed. To outline how the proposed methods could be applied for motion compensation, a feasible treatment scenario is described in this section.

The patient is first placed into an MR-scanner. Depending on the treatment approach, a breath-hold volume or 4DMRI data is gathered in an initial step. Based on this data, the population-based model is adapted to the patient's liver or a patient-specific motion model is created. In addition, the pathological tissue is marked as target and related to the motion model grid points. During sonication, the respiratory state is tracked by the MR-based 1D pencil beam. Based on these measurements, the future respiratory state of the liver is predicted and the position of the target is estimated by means of the motion model. This results in an interleaved acquisition of 1D pencil beams for tracking and MR-temperature maps to provide thermal dose feedback. The procedure is repeated until complete coagulation is guaranteed. Figure 2.10(a) shows the prototype of a MRgHIFU setup located at the University Hospital Geneva (HUG). The system is still under development. Figure 2.10(b) shows an example of a fast HIFU ablation of 120 seconds duration in a non-moving target. The white bars on the left and top correspond to 2 cm (1 cm/tick). The image shows a slice through the sonicated volume of size $2 \text{ cm} \times 2.5 \text{ cm} \times 2 \text{ cm} \approx 10 \text{ cm}^3$, which corresponds to a ablation rate of $\approx 5 \text{ cm}^3/\text{min}$.

Bibliography

- [1] WHO. Cancer, factsheet 297. 2012.
- [2] J W Hand, A Shaw, N Sathoo, S Rajagopal, R J Dickinson, and L R Gavrilov. A random phased array device for delivery of high intensity focused ultrasound. *Physics in Medicine and Biology*, 54(19):5675, 2009.
- [3] Vincent Auboiroux, Erik Dumont, Lorena Petrusca, Magalie Viallon, and Rares Salomir. An MR-compliant phased-array HIFU transducer with augmented steering range, dedicated to abdominal thermotherapy. *Physics in Medicine and Biology*, 56(12):3563, 2011.
- [4] Baudouin Denis Senneville, Mario Ries, Lambertus W. Bartels, and Chrit T. W. Moonen. Mri-guided high-intensity focused ultrasound sonication of liver and kidney. In Thomas Kahn and Harald Busse, editors, *Interventional Magnetic Resonance Imaging*, Medical Radiology, pages 349–366. Springer Berlin Heidelberg, 2012.
- [5] Francois Cornelis, Nicolas Grenier, Chrit T Moonen, and Bruno Quesson. In vivo characterization of tissue thermal properties of the kidney during local hyperthermia induced by mr-guided high-intensity focused ultrasound. *NMR in Biomedicine*, 24(7):799–806, 2011.
- [6] Derek L.G. Hill, Colin Studholme, and David J. Hawkes. Voxel similarity measures for automated image registration. *SPIE*, 2359:205–16, 1994.
- [7] T Rohlfing, CR Maurer, WG O’Dell, and JH Zhong. Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated MR images. *MEDICAL PHYSICS*, 31(3):427–432, March 2004.
- [8] Martin von Siebenthal, Gabor Székely, Urs Gamper, Peter Boesiger, Antony Lomax, and Ph. Cattin. 4D MR imaging of respiratory organ motion and its variability. *Physics in Medicine and Biology*, 52:1547–64, 2007.

- [9] Baudouin Denis de Senneville, Charles Mougenot, Bruno Quesson, Iulius Dragonu, Nicolas Grenier, and Chrit T. W. Moonen. Mr thermometry for monitoring tumor ablation. *European Journal of Radiology*, 17(9):2401–2410, 2007.
- [10] Baudouin Denis de Senneville, Mougenot C, and Moonen CTW. Real-time adaptive methods for treatment of mobile organs by MRI-controlled high-intensity focused ultrasound. *Magn. Reson. Med.*, 57(2):319–330, 2007.
- [11] Jan Lesniak, Tokuda J, Kikinis Ran, Burghart C, and Hata N. A device guidance method for organ motion compensation in MRI-guided therapy. *Phys. Med. Biol.*, 52(21):6427–6438, 2007.
- [12] M. Ries, B. D. Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen. Real-time 3d target tracking in mri guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine*, 64:01704–1712, 2010.
- [13] Mette K Stam, Sjoerd P M Crijns, Bernard A Zonnenberg, Maurits M Barendrecht, Marco van Vulpen, Jan J W Lagendijk, and Bas W Raaymakers. Navigators for motion detection during real-time mri-guided radiotherapy. *Physics in Medicine and Biology*, 57(21):6797, 2012.
- [14] Martin von Siebenthal, Gabor Székely, Antony Lomax, and Philippe Cattin. Inter-subject modelling of liver deformation during radiation therapy. In *MICCAI*, LNCS 4791, pages 659–66, 2007.
- [15] 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. *Transactions on Medical Imaging*, 18:712–21, 1999.
- [16] Martin von Siebenthal, Philippe Cattin, Urs Gamper, Antony Lomax, and Gábor Székely. 4D MR imaging using internal respiratory gating. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 3750, pages 336–43. Springer-Verlag GmbH, Oct 2005.
- [17] Patrik Arnold, Frank Preiswerk, Beat Fasel, Rares Salomir, Klaus Scheffler, and Philippe Cattin. Model-Based Respiratory Motion Compensation in MRgHIFU. *IPCAI*, LNAI 7330:54–63, 2012.
- [18] Patrik Arnold, Frank Preiswerk, Beat Fasel, Rares Salomir, Klaus Scheffler, and Philippe C. Cattin. 3d organ motion prediction for mr-guided high intensity focused ultrasound. In *MICCAI (2)*, pages 623–630, 2011.
- [19] Frank Preiswerk, Patrik Arnold, Beat Fasel, and Philippe C. Cattin. A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements. *Springer LNCS*, 7029:207–214, 2011.

- [20] Golnoosh Samei, Christine Tanner, and Gabor Székely. Predicting liver motion using exemplar models. In Hiroyuki Yoshida, David Hawkes, and Michael W. Vannier, editors, *Abdominal Imaging. Computational and Clinical Applications*, volume 7601 of *Lecture Notes in Computer Science*, pages 147–157. Springer Berlin Heidelberg, 2012.
- [21] Silja Kiriyanthan, Ketut Fundana, and Philippe C. Cattin. Discontinuity Preserving Registration of Abdominal MR Images with Apparent Sliding Organ Motion. In *Abdominal Imaging. Computational and Clinical Applications.*, volume 7029 of *LNCS*, pages 231–139. Springer-Verlag, 2012.

3 3D Organ Motion Prediction for MRgHIFU

Contents

The paper presented in this chapter describes a two-step prediction method dedicated to motion-compensated MRgHIFU treatments. The proposed method is a combination of a pattern matching approach using a subject-specific static motion atlas and a population-based statistical drift model. In a first step, the subject-specific 3D motion is acquired, processed and stored in an atlas. Local region tracking on the diaphragm is used to provide the respiratory signal which is used to index the atlas and to generate the predicted motion. In a second step, a population-based statistical drift model is used to compensate for organ deformation occurring over longer time-scales that have not been captured by the initially acquired atlas. After an atlas creation time of around 7-13 minutes, motion prediction experiments were evaluated for the subsequent 7-13 minutes. The application of the motion atlas alone results in a mean error of 1.6 mm. Combined with the statistical drift model, the error is reduced by an additional 30 % to 1.1 mm.

Outline

The research problem is introduced and motivated in Section 1. Section 2 presents the database, the creation of the static atlas and the prediction scheme based on this atlas. Furthermore, the population-based statistical drift model and its application are described. Section 3 presents the results of the experiments without and with drift compensation, before the paper concludes with a discussion and outlook in Section 4.

3D Organ Motion Prediction for MR-Guided High Intensity Focused Ultrasound

Patrik Arnold, Frank Preiswerk, Beat Fasel, Rares Salomir, Klaus Scheffler,
and Philippe C. Cattin

Medical Image Analysis Center, University of Basel, 4000 Basel, Switzerland
{patrik.arnold,philippe.cattin}@unibas.ch

Abstract. MR-guided High Intensity Focused Ultrasound is an emerging non-invasive technique capable of depositing sharply localised energy deep within the body, without affecting the surrounding tissues. This, however, implies exact knowledge of the target's position when treating mobile organs. In this paper we present an atlas-based prediction technique that trains an atlas from time-resolved 3D volumes using 4DMRI, capturing the full patient specific motion of the organ. Based on a breathing signal, the respiratory state of the organ is then tracked and used to predict the target's future position. To additionally compensate for the non-periodic slower organ drifts, the static motion atlas is combined with a population-based statistical exhalation drift model. The proposed method is validated on organ motion data of 12 healthy volunteers. Experiments estimating the future position of the entire liver result in an average prediction error of 1.1 mm over time intervals of up to 13 minutes.

1 Introduction

Respiratory organ motion is a complicating factor in the treatment of pathological tissue with MR-guided High Intensity Focused Ultrasound (MRgHIFU). Focused ultrasound has the unique capability to deposit sharply localised energy deep into the tissues, producing thermal ablation. Accurate spatial and rapid temporal beam spot focusing in the range of millimetres and within milliseconds, respectively, is reachable and hence increasing the demand of more exact knowledge about the organ's position. Accurate tracking of pathological tissue in mobile organs would not only increase patient safety, but also reduce the treatment time, as the gating window can be increased without sacrificing precision. Although the patient is located within the MR system during sonication, the scan-time is mainly required for the temperature feedback control of the HIFU system to determine the thermal dose given to a tumour. Non MR-based external surrogate markers would thus be ideal for the prediction of the respiratory-induced organ motion. Several techniques have been proposed in the literature to handle respiratory organ motion. Existing approaches that ensure comprehensive target coverage with minimal damage to the surrounding tissue include the optimisation of safety margins, voluntary or forced breath-hold, respiratory

gating or full tracking of the target. These methods were discussed in depth and compared in recent publications [1,2]. The holy grail of three-dimensional motion compensation in free-breathing awake patients is still out-of-reach, though. On the one hand, this goal could be reached by using ultrasound for real-time tracking [3]. On the other hand, Ries *et al.* [4] proposed only recently a real-time tracking method that observes the target on a 2D image plane combined with a perpendicular acquired pencil beam navigator, finally obtaining 3D information of the targets trajectories. The future target position is then estimated by a 3D Kalman filter. The method was tested in phantom experiments on human kidneys and in vivo with kidneys of ventilated pigs, both following a regular and stable breathing pattern. The tracking quality is evaluated by comparing temperature distribution obtained after 60s of HIFU application with and without motion compensation, resulting in higher final temperatures in the target area with enabled motion compensation.

In this paper, we present a novel atlas-based respiratory motion prediction method for free breathing patients. In contrast to the state-of-the-art, the slower modes of non-periodic organ deformation, that occur in addition to the fitful respiratory motion and that are not detectable by external sensors, are compensated by means of a population-based statistical drift model. Although the proposed generic framework is applicable to any abdominal organ, *e.g.* the kidney, the prediction technique is evaluated on real 4DMRI motion data of the liver.

2 Materials and Methods

2.1 Data Acquisition

To learn the patient setup specific breathing characteristics and organ motion, 4DMRI [5] sequences of 12 healthy volunteers (6 female, 6 male, average age 31, range 17-75) were acquired. During roughly one hour acquisition sessions, 14-26 minutes of time-resolved motion data was captured. MR volumes consisting of 25-30 slices covering the right liver lobe with a voxel size of $1.4 \times 1.4 \times 4 \text{ mm}^3$ and with a temporal resolution of 300-400 ms were obtained. The retrospectively reconstructed stacks cover the entire range of observed breathing depths. The vector fields describing the motion between the different respiratory states of the liver were estimated by means of 3D non-rigid registration [5,6] between the reconstructed volumes. In order to estimate the liver's future position we need to keep track of the current respiratory state on the basis of a breathing signal (surrogate marker). Regardless whether we measure the breathing signal by a breathing belt, by an optical chest wall tracker or by a pencil beam navigator placed on the diaphragm, the different respiratory states of the liver can reliably be tracked over a short period of time, as has been shown in [7]. In this work, we extracted a pencil beam navigator by tracking a manually defined region (Fig. 1(a)) on the navigator slice. The inferior-superior motion of the diaphragm was persistently tracked by template matching the dedicated region with all subsequent navigator frames throughout the acquisition sequence, providing one respiratory position per reconstructed volume (Fig. 1(b)).

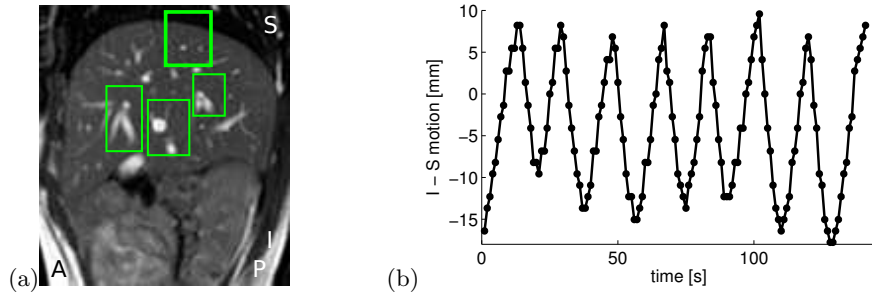


Fig. 1. (a) Sagittal view of a navigator slice with 4 marked regions used for slice stacking and drift compensation. The region indicated by the thicker frame is dedicated to track the diaphragm's position, providing the respiratory signal (b).

2.2 Atlas Creation

Using the 4D organ motion data of all 12 volunteers, we simulated a realistic MRgHIFU scenario. In particular, the first 7-13 minutes of 4DMRI scan time were used to build the motion atlas. This initial training time was long enough to cover all typical respiratory cycles. The remaining 4D motion data was used as ground truth to validate our prediction scheme. In order to keep correspondence between the acquisition of the atlas and the final treatment, the volunteers were asked not to move over the course of the entire sessions.

For each patient, a specific atlas and ground truth dataset is created, wherein both the breathing signal and corresponding 3D vector fields are stored pair-wise for each time step. An example of such an atlas is illustrated on the left side of Fig. 2 as well as the ground truth data for the validation on the right, both containing the respiratory signal and the associated organ displacement depicted by the black arrows between the reconstructed volumes. The breathing signal and the 3D vector fields describing the motion covering 150-400 breathing cycles or 1200-2000 time steps, respectively, serve as the atlas' database.

2.3 Motion Prediction

To readjust the treatment focus, any breathing-controlled tracking method must be able to estimate the target's position at some future time. This estimation must be based on measurements of the past breathing signal. Since our approach deals with a rather low sample rate, we use the atlas as combined breath and 3D motion look-up-table instead of an on-line learning based algorithm. However, only realistic, already seen motion patterns are being generated. Let S be the respiratory signal given as the sequence $S = s_j|_{j=1,\dots,m}$, with the indices denoting the running time steps t . At a given point in time j , the prediction provides an estimate $s_p = s'_{j+\Delta}$ of S and of the corresponding 3D motion vector field $\mathbf{u}_p = \mathbf{u}'_{j+\Delta}$, describing the future displacement of the organ for a later time point (Fig. 2). In the following experiments, we predicted $\Delta = 1$ time step into the future. One time step corresponds to roughly 300 ms, given by the

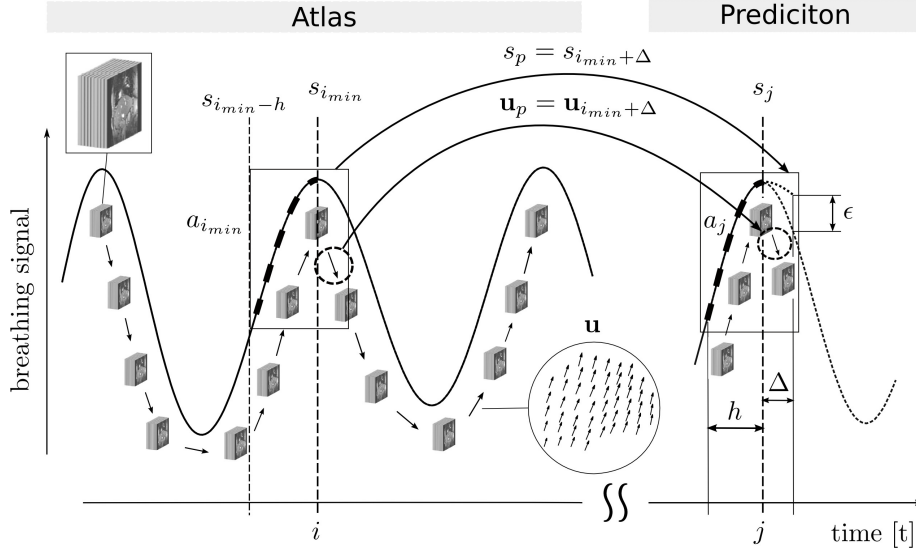


Fig. 2. Schematic illustration of the combined breath and motion atlas. Based on the signals history length h , the prediction yields the 3D displacement field \mathbf{u}_p that estimates the organ's future position, Δ time steps ahead.

4DMRI sequence. Note, that although the experiments are performed and validated on the temporal resolution of 3-4 Hz, Δ can be chosen arbitrary. In that case, the breathing signals and the vector fields are interpolated, allowing any predictive time gap and smooth beam re-focusing. For the prediction, we propose using the last h values of the breathing signals history denoted by the vector $\mathbf{a}_j = (s_{j-h}, \dots, s_j)$. The reference signal S_{ref} basically serving as the atlas is represented by similar vectors $\mathbf{a}_i = (s_{i-h}, \dots, s_i)$. The prediction of S at the time point j is chosen by finding the best match of the current breathing signal vector \mathbf{a}_j within the reference signals from the atlas:

$$i_{min} = \arg \min_i |\mathbf{a}_i - \mathbf{a}_j|. \quad (1)$$

The future run of $\mathbf{a}_{i_{min}}$ with minimum aberration from the history \mathbf{a}_j is chosen to estimate the organ's prospective respiratory state and corresponding displacement field, Δ time steps ahead:

$$s_p = s_{i_{min}+\Delta} \quad \text{and} \quad \mathbf{u}_p = \mathbf{u}_{i_{min}+\Delta}. \quad (2)$$

Finally, the task of predicting the organ's motion is handled by estimating the breathing signal's future evolution, yielding the well adapted displacement fields for the prediction. Since the algorithm is continuously adjusting to new input data, it can quickly adapt to the irregularity of the periods and amplitudes of the respiratory signal of a free breathing person.

2.4 Drift Compensation

Besides the displacements of the liver caused by respiratory motion, additional deformation independent of the fitful breathing motion can occur within a few minutes. Since the proposed prediction method base on collecting the patient-specific liver motion during the initial training phase of a couple of minutes, these organ drifts can not be captured during the short acquisition of the atlas. These drifts can quickly invalidate the applicability of the static atlas with the consequence of increasing systematic prediction errors. However, during sonication within the MR system, the functionality of the scanner is used for temperature feedback of the HIFU device and therefore, a scan time intensive 3D drift tracking is hard to achieve. On the other side, measuring a one dimensional breathing signal only, tracking the inferior-superior motion of the diaphragm respectively, is not sensitive to drifts in the inferior part of the liver (Fig. 3(a)). We propose to acquire one update-navigator slice after every 60s to capture the exhalation position of the liver based on the breathing signal. Comparing the displacements of the tracked regions with the regions on the actually acquired slice provides the needed information used for the correction of the previously acquired static atlas. In order to compensate these drifts, we introduce a population-based statistical drift model describing the inter-subject variations of exhalation positions in a shared shape-free coordinate system [8]. Shape-free means, that only the relative differences to the first exhalation position of each subject, the drifts, respectively, are used for modelling. Thereby, we assume that the drift is independent of the respiratory motion and is similar for all subjects. From each subject, 200 exhalation positions ($m = 11 \times 200$) with $N = 290$ corresponding points per liver ($n = 3N$), placed on a 3D regular grid with a 15 mm resolution, are mean-free concatenated in a data matrix $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) \in \mathbb{R}^{n \times m}$ with $\mathbf{x}_k = \mathbf{v}_k - \bar{\mathbf{v}}$ and sample mean $\bar{\mathbf{v}} = \frac{1}{m} \sum_{k=1}^m \mathbf{v}_k$. Applying Principal

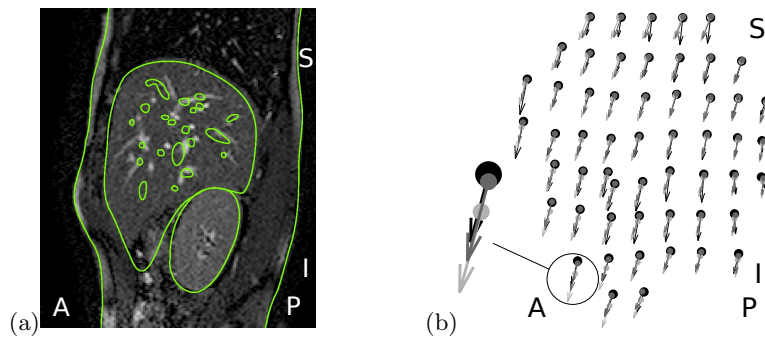


Fig. 3. (a) View of a sagittal placed navigator slice before (*edges*) and 20 minutes later after an exemplary drift displacement. The position of the diaphragm remains almost constant while the inferior part of the liver is drifting. (b) Ground truth motion field (*black*), prediction with static atlas (*light grey*) and with updated atlas (*dark grey*).

Component Analysis on the data, the vectors \mathbf{x} are defined by the coefficients c_k and the eigenvectors \mathbf{s}_k of $\mathbf{S} = (\mathbf{s}_1, \mathbf{s}_2, \dots)$ of the covariance matrix of the data:

$$\mathbf{x} = \sum_{k=1}^{m-1} c_k \sigma_k \mathbf{s}_k = \mathbf{S} \cdot \text{diag}(\sigma_k) \mathbf{c}. \quad (3)$$

Hereby, σ_k are the standard deviations within the data along each eigenvector \mathbf{s}_k . As elaborated in [9], the full vector \mathbf{x} can be found by an incomplete measurement $\mathbf{r} \in \mathbb{R}^l, l < n$ that minimises

$$E(\mathbf{x}) = \|\mathbf{L}\mathbf{x} - \mathbf{r}\|^2, \quad (4)$$

where \mathbf{L} represents a subspace mapping $\mathbf{L} : \mathbb{R}^n \mapsto \mathbb{R}^l$ such that $\mathbf{r} = \mathbf{L}\mathbf{x}$. The reduced version of \mathbf{S} can be written as $\mathbf{Q} = \mathbf{L}\mathbf{S} \cdot \text{diag}(\sigma_k) \in \mathbb{R}^{l \times m-1}$, yielding eigenvectors of the form $\mathbf{q}_k = \sigma_k \mathbf{L}\mathbf{s}_k \in \mathbb{R}^l$. The most probable organ deformation \mathbf{v} given the incomplete measurements \mathbf{r} is then

$$\mathbf{v} = \mathbf{S} \cdot \text{diag}(\sigma_k) \mathbf{c} + \bar{\mathbf{v}}, \text{ where } \mathbf{c} = \mathbf{Q}^+ \mathbf{r}. \quad (5)$$

Hereby \mathbf{Q}^+ is the pseudoinverse of \mathbf{Q} . The vector \mathbf{r} describes the relative differences from a few grid points at the beginning of the data acquisition to the actual exhalation position. These displacements are captured again by template matching the defined regions (Fig. 1(a)) with the update-navigator slice, measuring the distinct distances between the matching regions. Tracking 4 individual regions enables the detection of non-rigid deformations. As the centers of the templates may rarely coincide with the grid points of the model, the shifts of the templates have been adopted to the 3 closest points of the grid (12 out of 290 points), used as inputs for the drift model. The prediction from the static atlas \mathbf{u}_p (Eq. 2) is updated by the non-rigid correction field (Fig. 3(b)) provided by Eq. (5):

$$\hat{\mathbf{u}}_p = \mathbf{u}_p + \mathbf{v}. \quad (6)$$

3 Results

3.1 Motion Prediction without Drift Compensation

In a first approach we evaluated the capability of organ motion compensation by means of a static atlas without drift compensation. Based on the data mentioned in Sec. 2.2, motion prediction experiments were performed on 4DMRI datasets of all 12 volunteers. The parameter h introduced in Sec. 2.3 was optimised and found to work best for $h = 3$ time steps (≈ 1 s). The prediction experiments were evaluated for all subjects, covering 75-200 full respiratory breathing cycles. The error in prediction for $\Delta = 1$ time step (≈ 300 ms) was calculated point-wise over all grid points and time steps. The results are plotted in Fig. 4(a), characterised by the median, 5th and 95th percentiles. The dashed line is set to 2 mm, marking an acceptable precision limit for HIFU treatments [10]. The impact of the liver's

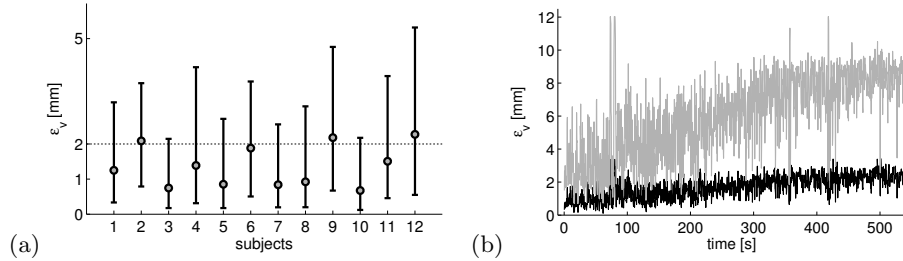


Fig. 4. (a) Resulting deviations between predicted and ground truth liver motions for 12 different subject over time intervals up to 13 minutes. Error bars around the median show the 5th and 95th percentile deviation. (b) Mean (*black*) and maximum error (*grey*) of motion prediction based on the static atlas for the drifting liver of subject 4.

drift is clearly visible in Fig. 4(b), when monitoring the prediction performance over several minutes. The average error over all subjects is 1.6 mm.

3.2 Motion Prediction with Drift Compensation

With equal settings as in Sec. 3.1, the same experiments but with drift compensation as elaborated in Sec. 2.4, were realised (Fig. 5(a)). The statistical drift models were built from 11 of 12 livers in leave-one-out experiments. Although the residual MR time during HIFU treatment is rather sparse, we allowed the acquisition of one 2D navigator slice every 60s, capturing the actual exhalation position. This time interval is based on the maximal observed drifting speed of 0.5 mm/min. Following Eq. (5) and (6), the most probable drift deformation of the left-out liver is provided by the model and used as a drift-update of the previously acquired static atlas. Taking the drift into account, the prediction performance remains constant over time as shown in Fig. 5(b). The error averaged over all subjects improved by 30% to 1.1 mm, with a notable impact for

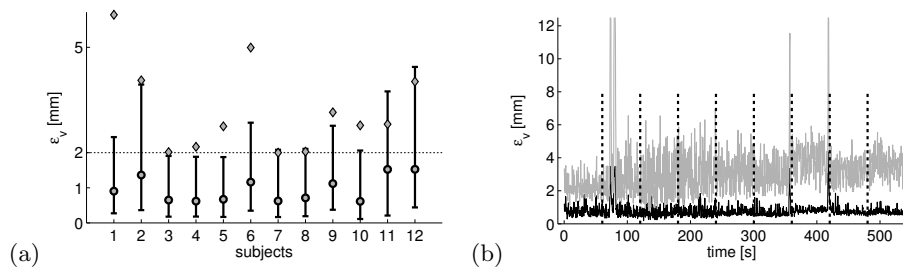


Fig. 5. (a) Residual error of motion prediction with drift compensation and median error without any motion compensation (\diamond). (b) Mean (*black*) and maximum error (*grey*) of the prediction with drift compensation every 60 s (*black-dotted lines*).

the subjects 4,9 and 12. Without any motion compensation, the mean prediction error would be 4.7 mm.

4 Discussion and Outlook

Despite frequently occurring organ drifts, our proposed method proved to be reliable enough for the application in MRgHIFU systems. Using the R software package (Version 2.11.1), we used the Kolmogorof-Smirnov test to test the mean errors of both experiments for normality. Assuming a significance level of 0.05, the t-tests showed that the statistical drift model significantly improved the prediction accuracy ($p < 0.05$). By replacing the pencil beam navigator with a faster low lag signal, such as the breathing belt or a spirometer, the prediction quality should further improve, as the time span the system has to predict into the future decreases.

Acknowledgments. This work has been supported by the research network of the Swiss National Science Foundation-Project Nr. CR32I3_125499.

References

1. Keall, P., Mageras, G., Balter, J., Emery, R., Forster, K., Jliang, S., Kapatoes, J., Low, D., Murphy, M., Ramsey, B.M.C., Herk, M.V., Vedam, S., Wong, J., Yorke, E.: The management of respiratory motion in radiation oncology report of AAPM task group. *Med. Phys.* 33, 3874–3900 (2006)
2. Webb, S.: Motion effects in (intensity modulated) radiation therapy. *Phys. Med. Biol.* 51, R403–R425 (2006)
3. Pernot, M., Tanter, M., Fink, M.: 3-D real-time motion correction in high-intensity focused ultrasound therapy. *Ultrasound in Medicine and Biology* 30(9), 1239–1249 (2004)
4. Ries, M., de Senneville, B.D., Roujol, S., Berber, Y., Quesson, B., Moonen, C.: Real-time 3D target tracking in mri guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine* 64, 1704–1712 (2010)
5. von Siebenthal, M., Szekely, G., Gamper, U., Boesiger, P., Lomax, A., Cattin, P.: 4D MR imaging of respiratory organ motion and its variability. *Phys. in Med. Biol.* 52, 1547–1564 (2007)
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. *Transactions on Medical Imaging* 18, 712–721 (1999)
7. von Siebenthal, M., Szekely, G., Lomax, A.J., Cattin, P.C.: Systematic errors in respiratory gating due to intrafraction deformations of the liver. *Med. Phys.* 34, 3620–3629 (2007)
8. von Siebenthal, M., Székely, G., Lomax, A., Cattin, P.C.: 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)
9. Blanz, V., Vetter, T.: Reconstructing the complete 3D shape of faces from partial information. *Informationstechnik und Technische Informatik* 44, 295–302 (2002)
10. Vedam, S., Kini, V.R., Keall, P.J., Ramakrishnan, V., Mostafavi, H., Mohan, R.: Quantifying the predictability of diaphragm motion during respiration with a non-invasive external marker. *Med. Phys.* 30, 505–513 (2003)

4 Model-Based Respiratory Motion Compensation in MRgHIFU

Contents The paper described in this chapter introduces a respiratory motion compensation method consisting of a subject-specific breathing model [18] and a population-based statistical motion model [19]. A lightweight respiratory breathing model, capturing the patient-specific breathing characteristics, is acquired and used to predict future respiratory states of the liver. Based on this prediction, a population-based motion model is applied to estimate the displacement of the entire liver. The use of a population-based model has the advantage that no patient-specific 3D motion data needs to be acquired and processed. Although no patient-specific motion model was created, a prediction performance of 1.7 mm was obtained.

Based on Arnold P., Preiswerk F., Fasel B., Salomir R., Scheffler K., and Cattin, P. C. (2011). 3D Organ Motion Prediction for MR-Guided High Intensity Focused Ultrasound. In MICCAI, LNCS 6892, page 623-630.

Preiswerk F., Arnold P., Fasel B., and Cattin, P. C. (2011). A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements. Springer LNCS 7029, page 207-214.

Outline After the introduction of the research problem in Section 1, Section 2 presents the database, the temporal prediction as well as the usage of the statistical motion model for motion prediction. Section 3 presents the results for both the 1D breath and the motion prediction. Section 4 concludes with a discussion of the results and an outlook for future work.

Model-Based Respiratory Motion Compensation in MRgHIFU

Patrik Arnold¹, Frank Preiswerk¹, Beat Fasel¹, Rares Salomir², Klaus Scheffler³,
and Philippe Cattin¹

¹ Medical Image Analysis Center, University of Basel, Switzerland
patrik.arnold@unibas.ch

² Radiology Department, University Hospitals of Geneva, Switzerland

³ Department of Neuroimaging, University of Tuebingen, Germany

Abstract. Magnetic Resonance guided High Intensity Focused Ultrasound (MRgHIFU) is an emerging non-invasive technology for the treatment of pathological tissue. The possibility of depositing sharply localised energy deep within the body without affecting the surrounding tissue requires the exact knowledge of the target's position. The cyclic respiratory organ motion renders targeting challenging, as the treatment focus has to be continuously adapted according to the current target's displacement in 3D space. In this paper, a combination of a patient-specific dynamic breath model and a population-based statistical motion model is used to compensate for the respiratory induced organ motion. The application of a population based statistical motion model replaces the acquisition of a patient-specific 3D motion model, nevertheless allowing for precise motion compensation.

1 Introduction

Focused Ultrasound deposits sharply localised energy in the tissue causing thermal ablation. Precise targeting demands for exact knowledge of the target's position. The compensation of the fitful respiratory organ motion is a challenging task in the treatment of pathological tissue in abdominal organs. If breathing motion is not compensated, the exposure of healthy tissue increases and the thermal dose delivered to the tumour is reduced. Continuous target displacement tracking in 3D space requires accurate spatial and rapid temporal beam refocusing in the range of millimetres and milliseconds, respectively. Any realisation of a real-time target tracking-based dose delivery must thus be able to predict the target's position at some future time in order to compensate for the finite time delay between the acquisition of the current target's position and the mechanical response of the system to change treatment focus.

During sonication the Magnetic Resonance (MR) scan-time is mainly required for the temperature feedback control of the High Intensity Focused Ultrasound (HIFU) system, quantifying the thermal dose given to the tissue in order to guarantee complete coagulation of the tumour. Therefore, not enough MR scan-time is left to track the tumour in 3D. To determine the thermal dose, temperature maps in regular distances around the tumour are acquired. Similarly as proposed in [1], the navigator (pencil beam) feedback information is used to reposition the temperature mapping slice to resolve organ

displacements. In this work, we propose to use this 1-dimensional navigator feedback information not only to track the current respiratory state, but also to predict the organ's future displacement, e.g. the position of the tumour.

Several approaches have been proposed to track and predict the motion of abdominal organs. Ries *et al.* [1] proposed a real-time tracking method that observes the target on a 2D image plane combined with a perpendicular acquired pencil beam navigator, providing quasi-3D information of the target trajectories. The future 3D target position is then estimated by a Kalman filter. Underlying a regular and stable breathing pattern, the method was tested in phantom experiments and in vivo on ventilated pigs. The accuracy of the approach is not evaluated on ground truth motion data, but by indirectly comparing the temperature maps obtained after 60 seconds of HIFU sonication with and without motion compensation, resulting in higher maximal temperatures in the target area with enabled motion compensation. However, the experiments have neither been evaluated on ground truth data nor under free breathing conditions.

Ruan and Keall [2] proposed a predictor based on Kernel Density Estimation to account for system latencies caused by software and hardware processing. They use 3D motion trajectories of implanted markers to train the predictor in a lower dimensional feature space using Principal Component Analysis (PCA). The prediction is performed in this subspace and mapped back into the original space for the evaluation. The drawback of the method is that only the position of directly observed internal fiducials can be predicted and not of the entire organ.

Only recently, a combination of a pattern matching approach using a static subject-specific model and a population-based statistical drift model for motion-compensated MRgHIFU treatment was described and evaluated on realistic 4DMRI data [3]. While the results are convincing, the acquisition of a patient-specific 3D motion atlas takes several minutes and the processing time is in the range of hours and thus is not acceptable for clinical use. In particular, the multiple volume-to-volume registrations take up to several hours, in which the patient is asked not to move in order to stay aligned with the acquired model.

Preiswerk *et al.* [4] showed, that the displacement of the entire liver can be spatially predicted by tracking three well distributed markers (implanted fiducials) within the liver using a population-based statistical motion model. Based on an exhalation breath-hold scan, accurate prediction is achieved. Dispensing with the need of extensive pretreatment volume imaging and its time consuming 3D non-rigid registration, no attention is paid to a potential system lag, which is essential for real-time tracking. Also this method is based on full 3D motion information of implanted markers.

The main contribution of the presented work is the combination of a patient-specific fast and lightweight respiratory breathing model and a population-based motion model to a novel, completely non-invasive and clinically feasible 3D motion compensation method for MRgHIFU treatments. The proposed method addresses certain weaknesses of the state-of-the-art methods in terms of real-time usage and validation. On the one hand, the completely MR-based respiratory signal is continuously acquired and used to predict the organs future respiratory state in order to bridge the system's time delay between the tracking and treatment of the target. On the other hand, the

population-based motion model is applied to estimate the motion of the unobserved liver, without the need of acquiring a subject-specific 3D motion model.

2 Materials and Methods

For the evaluation of our approach, a realistic MRgHIFU scenario was assumed. During HIFU sonication, the measured information of the pencil beam navigator, *i.e.* the inferior-superior displacement (1D) of the diaphragm, is used as the *breathing signal*. Based on this *breathing signal* a patient-specific respiratory model is created, whereby a temporal prediction of the diaphragm's future position is estimated (Sec. 2.2). Having an estimate of this displacement, the population-based statistical model is used to compute the most likely 3D displacement of the entire liver, further referred as to *reconstruction* (Sec. 2.3).

2.1 Data and Ground Truth

The ground truth data was acquired by 4DMRI, a dynamic 2D MR imaging method capturing the respiratory motion during free breathing [5]. Thanks to the sagittal slice orientation and the interleaved acquisition of data slices and a dedicated so-called *navigator slice* at a fixed position, vascular structures used for the 3D reconstruction of the volumes are visible during complete breathing cycles and can be tracked with minimal out-of-plane motion. 4DMRI sequences of 20 healthy volunteers (mixed sexes, age range: 17-75) were captured. During acquisition sessions of roughly two hours, 20-45 minutes of time-resolved organ motion data was measured. MR volumes consisting of 25-30 slices (120×192 pixel) covering the right liver lobe with a voxel size of $1.4 \times 1.4 \times 4$ mm³ and with a temporal resolution of 300-400 ms were obtained. The retrospectively reconstructed 3D stacks cover the entire range of observed breathing depths. By means of B-spline-based 3D non-rigid registration [6], dense spatio-temporal vector fields describing the motion between the different respiratory states of the liver are extracted. The first manually segmented liver exhalation stack is taken as reference volume upon which the subsequent 3D stacks are incrementally registered from time-step to time-step. The vector field from the previous step is taken as an initial estimation, significantly speeding up the registration time and making the registration more robust by reducing the chance of getting trapped in a local minima. The resulting vector fields, describing the liver's displacements relative to the reference volume, serve as the basic data for the motion model and its evaluation in cross-validation experiments.

In order to build a statistical model from this data, inter-subject correspondence had to be established. For each subject mechanical corresponding points were manually selected on the reference volume surfaces in order to align the 20 datasets. These points mark the delineations between the superior surface in contact with lung, the anterior and the posterior areas, which slide along the abdominal wall, and the inferior surface. An isotropic grid with 10 mm resolution was placed in the resulting average liver and then transformed to the shape of each of the subjects. This finally gave a set of 20 topologically equivalent 3D liver volumes as well as vector fields describing the motion

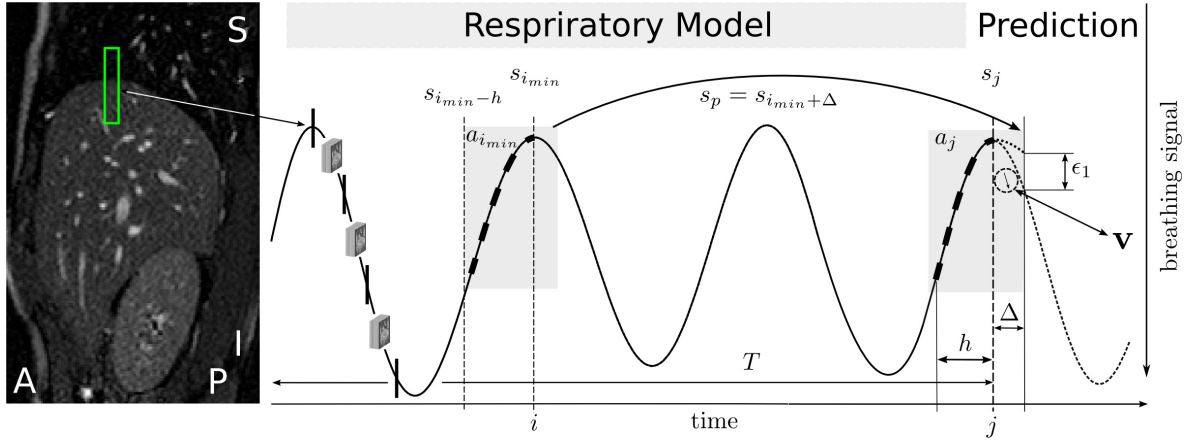


Fig. 1. Schematic illustration of combined respiratory model and motion model-based prediction. Based on the respiratory signal (|) captured at the marked diaphragm region, the displacement s_p of the diaphragm is predicted and from that the full liver displacement \mathbf{v} is reconstructed.

for each of the $N = 1261$ inter-subject corresponding grid points. For more detailed information we refer the reader to the article of Preiswerk *et al.* [4].

In this work, the described *breathing signal* is generated by simulating a pencil beam navigator placed on the acquired *navigator slices*. A manually defined region placed anywhere at the diaphragm was persistently tracked by template matching (Normalised Cross Correlation) throughout the acquisition sequence providing one respiratory position and displacement per acquired *navigator slice*, respectively. The inferior-superior component of the templates motion is interpreted as the *breathing signal* as obtained by a common pencil beam navigator, see Figures 1 and 2(a). The spatial resolution is thus given by the image's pixel size of roughly $1.4 \times 1.4 \text{ mm}^2$.

Since the 3D volumes are reconstructed at the time point between two *navigator slices* (see Figure 1, *left*) we linearly interpolate the *breathing signal* in order to obtain the respiratory positions and the 3D volumes at the same time points for the evaluation. In the following we deal with a linearly interpolated *breathing signal* with a sample rate of 6-8 Hz.

2.2 Temporal Prediction

Figure 1 schematically illustrates the prediction scene for the combined patient-specific and population-based model. As described above, the *breathing signal* is extracted by tracking a defined region on all the *navigator slices* followed by linear interpolation obtaining the intermediate respiratory states, where the ground truth 3D data is available for the validation.

The temporal prediction of the *breathing signal* is necessary in order to compensate for the system lag, caused by the pencil beam acquisition time, the processing of the data and the time for refocusing the HIFU beam to the newly calculated target. Any breathing-controlled tracking method must thus be able to estimate the target's position at some future time. The prediction of the future curve of the *breathing signal* is a key part of the prediction pipeline. Faulty predictions lead to wrong assumptions on

the diaphragm's displacement and thus to wrong spatial *reconstructions* of the whole liver displacement. Since the breathing pattern of a free breathing patient is very irregular over time, e.g. the amplitude and phase are changing nearly unpredictably, we use a prediction algorithm which can quickly adapt to the new input data. In the proposed method, however, the tracking of the respiratory state during sonication is based on pencil beams, therefore, one can expect a much lower sampling rate, as for example given by an optical tracking system. In our simulation we deal with a sampling rate of 6-8 Hz. Due to the low sampling rate, the learning based algorithms would lead to considerable prediction errors at each ex- and inhalation position before adapting. Therefore we use a similar technique as proposed in [3], where a one-dimensional breathing model based on the measured pencil beam navigators is created. In contrast to the latter approach where the model is acquired in a training phase and then stays fixed, our respiratory model steadily grows even during increasing treatment time T . Each newly measured data point (pencil beam position) is added to the model, thus getting more and more stable over time. As the prediction algorithm prefers the most recent measurements in the model, the model can be kept small to avoid a system slowdown caused by the increasing model size. All the data stored in the model is observed for the patient-specific operational setup, therefore only realistic displacements of the liver are predicted. For anomalous breathing patterns with a deviation from the breath model above a certain threshold, *i.e.* no matching pattern is found (*e.g.* coughing, new pattern), the HIFU beam can be switched off to ensure patient safety.

The model is best represented by a matrix \mathbf{A} , wherein the *breathing signal* is piecewise stored:

$$\mathbf{A} = \left(\begin{array}{cccc|c} s_1 & s_2 & \dots & s_h & s_{h+\Delta} \\ s_2 & s_3 & \dots & s_{h+1} & s_{h+1+\Delta} \\ \vdots & \vdots & & \vdots & \vdots \\ s_{i-h} & s_{i-h+1} & \dots & s_i & s_{i+\Delta} \\ \vdots & \vdots & & \vdots & \vdots \\ s_{T-h} & s_{T-h+1} & \dots & s_T & s_{T+\Delta} \end{array} \right). \quad (1)$$

The temporal prediction is based on the last h values of the current *breathing signal* given by the vector $\mathbf{a}_j = (s_{j-h}, \dots, s_j)$, where the index j denotes the actual time point. The prediction provides an estimate $s_p = s'_{j+\Delta}$ describing the future signal curve for a later time point, Δ time steps ahead. The best matching pattern of the current *breathing signal* vector \mathbf{a}_j and the column vectors \mathbf{a}_i of \mathbf{A} , is found with:

$$i_{min} = \arg \min_i \{ |\mathbf{a}_i - \mathbf{a}_j|, |j - T| \}. \quad (2)$$

The future curve of $\mathbf{a}_{i_{min}}$ with minimum aberration from the actual signal's history \mathbf{a}_j is considered as best estimate of the organ's future respiratory state:

$$s_p = s_{i_{min}+\Delta}. \quad (3)$$

The resulting prediction error ϵ_1 is then given by:

$$\epsilon_1 = |s_{j+\Delta} - s_p|. \quad (4)$$

The value s_p is the predicted shift in inferior-superior direction of the next diaphragm position. This displacement serves as the input to the motion model that then predicts the position of the entire liver. As the algorithm is continuously adjusting to new input data and updated with the new measured signal input, it can quickly adapt to the irregularity of the periods and amplitudes of the respiratory signal of a free breathing person. Figure 2(b) shows 60 seconds of robust 170 ms ahead prediction performance of an irregular breathing pattern measured by template matching (*blue*) and the model-based prediction (*green*) of subject 4.

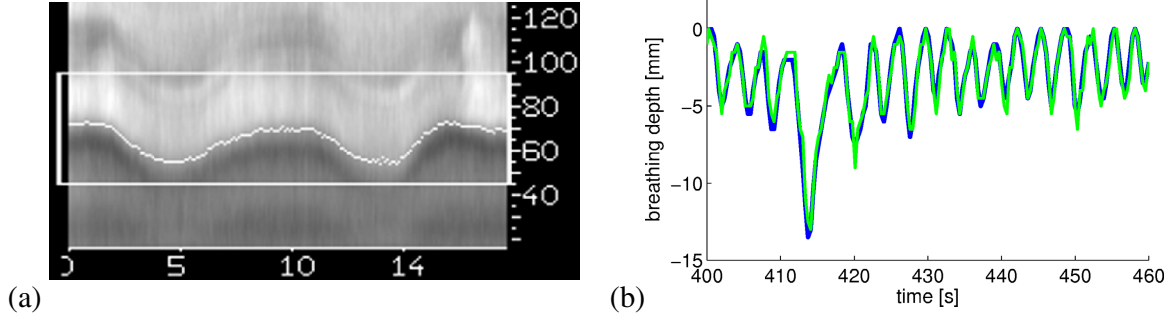


Fig. 2. (a) Typical pencil-beam navigator for MR thermometry real-time slice correction acquired at 10 Hz. (b) Example of 170 ms ahead prediction of an irregular breathing pattern of subject 4. Blue: tracked *breathing signal*; Green: robust respiratory model-based prediction.

2.3 Statistical Modelling

So far, the displacement of only one single point at the diaphragm is known from the prediction. The observed region, the centre of the pencil beam navigator template located on the *navigator slice*, respectively, has to be adopted to the closest grid-point of the subject's liver. The predicted shift s_p is then rigidly assigned to the corresponding model grid-point and the population-based statistical model is used for the *reconstruction* of the entire non-rigid liver displacement. From each of the 20 subjects, the vector fields of the first 15 breathing cycles are taken to build the model. The liver displacements are represented by a $3N$ -dimensional vector $\mathbf{v} = (\Delta u_1, \Delta v_1, \Delta w_1, \dots, \Delta u_N, \Delta v_N, \Delta w_N)^T$. Note, that the difference vector \mathbf{v} contains no shape information, but only the relative displacements with respect to the reference volume. The vector fields are mean-free concatenated in a data matrix $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) \in \mathbb{R}^{3N \times m}$ with $\mathbf{x}_k = \mathbf{v}_k - \bar{\mathbf{v}}$ and sample mean $\bar{\mathbf{v}} = \frac{1}{m} \sum_{k=1}^m \mathbf{v}_k$. Applying PCA to the data, the vectors \mathbf{x} are defined by the coefficients c_k and the Eigenvectors \mathbf{s}_k of $\mathbf{S} = (\mathbf{s}_1, \mathbf{s}_2, \dots)$ of the covariance matrix of the data:

$$\mathbf{x} = \sum_{k=1}^{m-1} c_k \sigma_k \mathbf{s}_k = \mathbf{S} \cdot \text{diag}(\sigma_k) \mathbf{c}. \quad (5)$$

Hereby, σ_k are the standard deviations within the data along each eigenvector \mathbf{s}_k . As elaborated in [7], the model coefficient \mathbf{c} for the full vector \mathbf{x} can be found by an incomplete estimate $s_p \in \mathbb{R}^l, l < N$ that minimises

$$E(\mathbf{c}) = \|\mathbf{Q}\mathbf{c} - s_p\| + \eta \cdot \|\mathbf{c}\|^2, \quad (6)$$

with $\mathbf{Q} = \mathbf{L}\mathbf{S} \cdot \text{diag}(\sigma_k)$, where \mathbf{L} represents a subspace mapping $\mathbf{L} : \mathbb{R}^N \mapsto \mathbb{R}^l$. In the case of a noisy or incorrect assumption s_p , tuning the regularisation factor η allows for *reconstructions* closer to the average quantified by the Mahalanobis distance $\|\mathbf{c}\|^2$. Solving Eq. 6 for \mathbf{c} with the singular value decomposition of $\mathbf{Q} = \overline{\mathbf{V}}\mathbf{W}\mathbf{V}^T$, yields:

$$\mathbf{c} = \mathbf{V} \text{diag}\left(\frac{w_k}{w_k^2 + \eta}\right) \overline{\mathbf{V}}^T s_p. \quad (7)$$

Using Eq. 7 the most probable organ displacement under the constraint of the known one-dimensional point-shift prediction s_p is then given by:

$$\mathbf{v} = \mathbf{S} \cdot \text{diag}(\sigma_k) \mathbf{c} + \bar{\mathbf{v}}. \quad (8)$$

The elaborated framework allows to associate the rigid 1D shift of 1 point placed at the diaphragm with the non-rigid 3D motion of the entire liver based on population statistics.

3 Experiments and Results

To evaluate the prediction performance of the algorithm for clinical relevant motion compensation, experiments on 20 volunteer subjects were performed. On average, displacements of the diaphragm from 5.5 mm to 15.2 mm in inferior-superior direction depending on the subject were observed. For simplicity of generating population statistics, the same amount of data from each subject was included for the experiments. For each experiment 1500 time steps, corresponding to 7-11 minutes, have been predicted.

In a first step, the prediction performance of the respiratory model is tested and evaluated on each of the subjects. In a second step, the respiratory model and the motion model prediction are evaluated in combination with cross-validation experiments. The predictive scene was evaluated every 300-400 ms, at the time points where the ground truth 3D data is available. All experiments were performed with a lookahead length of $\Delta = 1$, *i.e.* 150-200 ms and based on a signal history length of $h = 4$, corresponding to roughly 0.7 s.

3.1 Breath Prediction

Theoretically, the algorithm is able to predict after the first $h = 4$ time steps (≈ 0.7 s). But as more breathing cycles are collected in the respiratory model the more robust the method is predicting. Therefore, we observed the behavior of predictive performance as a function of time, *i.e.* with an increasing model size. Figure 3(a) shows the average error cumulated up to the given time on the axis and error bars showing the standard deviation. The error in prediction is retrospectively computed according to Eq.(4).

In Figure 3(b) the overall results of breath prediction for all 20 subject are visualised by error bars, marking the average and standard deviations. The experiments are evaluated after a model acquisition time of 60 seconds. The average error over all subjects is 0.6 mm with an observed average breathing depth of 8.4 mm.

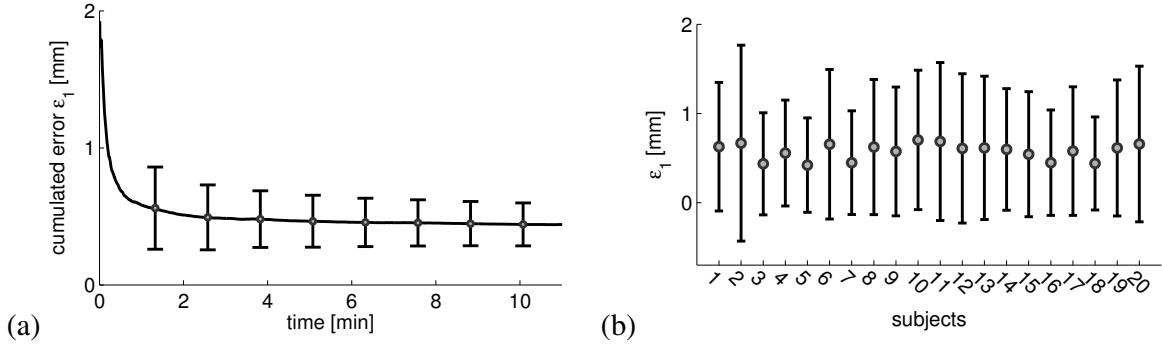


Fig. 3. (a) Average and standard deviations of the cumulated breath prediction error averaged over all subjects. The performance stabilises after a few minutes and is acceptable after 60 seconds. (b) Prediction performance of the respiratory model evaluated for 20 subjects averaged over a prediction length of 7-11 minutes, whereby the first 60 seconds are used to acquire a minimal model. The results are presented by error bars, marking the average and standard deviations for lookahead time of $\Delta_t \approx 180$ ms and signal history length of $h_t \approx 0.7$ s with an overall error of 0.6 mm.

3.2 Motion Model Prediction

The minimum size of population data to create a reliable model is still a unsolved problem as the exact distribution of the data for the entire population is unknown. The suitability of statistical models can, however, be shown empirically in cross-validation experiments. For the evaluation of our motion prediction technique leave-one-out statistical models of all the 20 subjects were computed. From the left-out data a respiratory signal was generated and used as test signal. As explained in Section 2.2 and 2.3, the respiratory motion of the full liver is predicted from one single point at the diaphragm only. For the *reconstruction* we took the 9 first principal components ending up with a model covering 98% of the variance of the original motion data. For each subject the manual segmentation of a reference volume and establishing correspondence (Sec. 2.1) is necessary.

As the predicted shift s_p can not fully be accounted for, the regularisation factor of Eq. (7) was set to $\eta = 5.5$ in order to get more plausible *reconstructions*. The error of prediction is determined by the point-wise Euclidean distance from the predicted liver motion to the ground truth motion of the left-out liver. To give an overview of the error distribution the results are visualised in Figure 4(a) by the median and error bars marking the 25th and 75th percentiles. The dashed line is set to 2 mm, marking an acceptable accuracy limit for HIFU treatments [8]. The average error over all subjects is 1.7 mm, in contrast to the average error without any motion compensation of 3.8 mm. In the case of no motion compensation, the error equals to the mean of the Euclidean distances to the reference volume over time. The spatial distribution of the averaged error over all subjects and time steps is shown in Figure 4(b). The root cause of the error are false predictions in inferior-superior and anterior-posterior direction with a maximal error of 2 mm, 1.1 mm and a minor error in left-right direction of 0.4 mm, respectively.

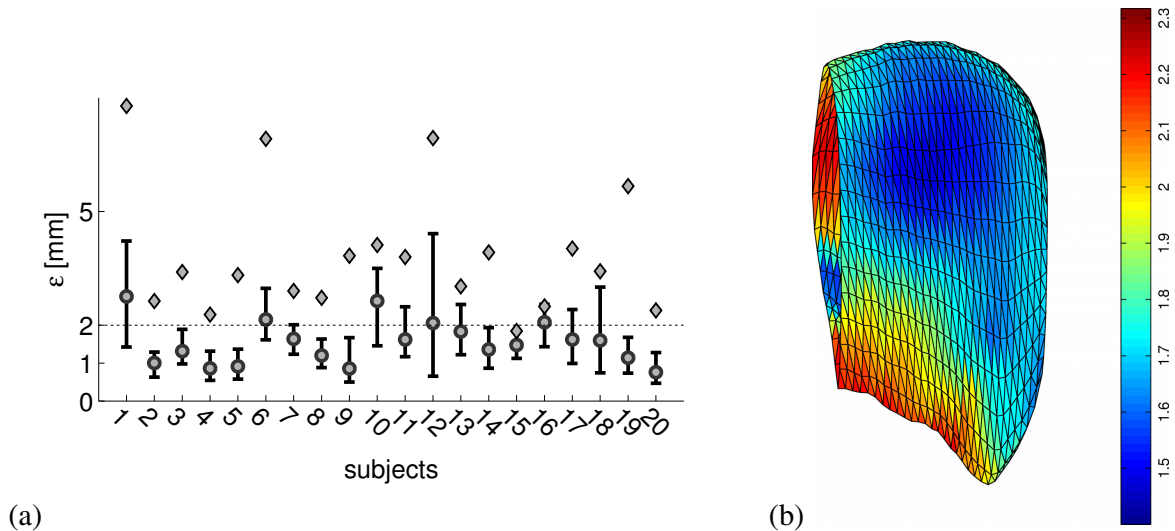


Fig. 4. (a) Resulting deviations between predicted and ground truth liver motions for 20 different subject over a time interval of 7-11 minutes. Error bars around the median show the 25th and 75th percentile deviation and mean error without any motion compensation (\diamond). (b) Averaged liver surface from 20 subjects of the right liver lobe at exhalation in anterior view. The colors represent the motion prediction error (in mm) averaged over 20 subjects at the liver's surface.

4 Conclusion

We presented a completely non-invasive and purely MR-based tracking method to predict the liver's 3D motion in real-time under free breathing. The method is a combination of a pattern matching approach to predict the patient-specific breathing pattern and a population-based statistical motion model based on PCA to *reconstruct* the respiratory induced organ motion. In the presented work, we demonstrate a safe and efficient technique for MRgHIFU treatment of pathological tissue in moving organs. Although the prediction technique is evaluated on real 4DMRI motion data of the liver, the proposed generic framework is applicable to any abdominal organ, *e.g.* the kidney. The method is evaluated on 4DMRI datasets of 20 healthy volunteers achieving an overall prediction error of 1.7 mm, where the predictive method is clinically applicable after 60 seconds.

Although the overall prediction error of our novel method is slightly higher than the state-of-the-art methods, the proposed technique addresses important issues for the non-invasive real-time application of MRgHIFU treatment in moving abdominal organs. Preiswerk *et al.* [4] achieve a prediction error of 1.2 mm by accurately knowing the 3D displacements of three well distributed points within the liver (*e.g.* implanted surrogate markers). In [3] a prediction error of 1.1 mm is achieved by acquiring 3D information of the patient specific liver motion.

In this work, however, a non-invasive MR-based tracking method is used, allowing to measure the 1-dimensional displacement of a single point on the diaphragm only. We are fully aware of that the second order organ deformation occurring over large time scales, the so called drifts, are not detectable by measuring a single point at the diaphragm only. But, since we are predicting over a short period of time, the different

respiratory states of the liver can reliably be tracked, as has been shown in [9]. Besides a 3D exhalation breath-hold scan, no patient-specific 3D motion data has to be acquired and processed in a pretreatment phase.

In future work we will investigate the possibility of better adapting the population-based motion model to a specific subject. Using a fast MR acquisition sequence, we plan on better restricting the population-based statistical motion model to a specific patient.

Acknowledgments. This work has been supported by the research network of the Swiss National Science Foundation-Project Nr. CR32I3_125499.

References

1. Ries, M., de Senneville, B.D., Roujol, S., Berber, Y., Quesson, B., Moonen, C.: Real-Time 3D Target Tracking in MRI Guided Focused Ultrasound Ablations in Moving Tissues. *Magnetic Resonance in Medicine* 64, 1704–1712 (2010)
2. Ruan, D., Keall, P.: Online Prediction of Respiratory Motion: Multidimensional Processing with Low-Dimensional Feature Learning. *Physics in Medicine and Biology* 55, 3011–3025 (2010)
3. Arnold, P., Preiswerk, F., Fasel, B., Salomir, R., Scheffler, K., Cattin, P.C.: 3D Organ Motion Prediction for MR-Guided High Intensity Focused Ultrasound. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part II. LNCS, vol. 6892, pp. 623–630. Springer, Heidelberg (2011)
4. Preiswerk, F., Arnold, P., Fasel, B., Cattin, P.C.: A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) Abdominal Imaging. LNCS, vol. 7029, pp. 207–214. Springer, Heidelberg (2012)
5. von Siebenthal, M., Székely, G., Gamper, U., Boesiger, P., Lomax, A., Cattin, P.: 4D MR Imaging of Respiratory Organ Motion and its Variability. *Phys. in Med. Biol.* 52, 1547–1564 (2007)
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. *Transactions on Medical Imaging* 18, 712–721 (1999)
7. Blanz, V., Vetter, T.: Reconstructing the Complete 3D Shape of Faces from Partial Information. *Informationstechnik und Technische Informatik* 44, 295–302 (2002)
8. Vedam, S., Kini, V.R., Keall, P.J., Ramakrishnan, V., Mostafavi, H., Mohan, R.: Quantifying the Predictability of Diaphragm Motion During Respiration with a Noninvasive External Marker. *Med. Phys.* 30, 505–513 (2003)
9. von Siebenthal, M., Székely, G., Lomax, A.J., Cattin, P.C.: Systematic Errors in Respiratory Gating due to Intrafraction Deformations of the Liver. *Med. Phys.* 34, 3620–3629 (2007)

5 Fast Manifold-based 4DMRI for Patient-Specific Organ Motion Modelling in MRgHIFU

Contents

This chapter's paper describes the fast acquisition of patient-specific 3D motion data and the efficient construction of patient-specific motion models. As the models are acquired immediately before the treatment, the method is applicable to any organ, patient pose and treatment setup. This addresses the lack of generality of population-based models, which are made for a specific organ and treatment setup. To decrease the pre-operative data acquisition time, the minimal amount of required 4DMRI data for motion modelling is investigated and novel 3D volume reconstruction techniques based on manifold learners are proposed. The goal of the work is to acquire patient-specific motion models in a short time without sacrificing precision.

Outline

Section 1 introduces the problem statement by summarising the state-of-the-art motion modelling approaches. Section 2 presents the database and describes the tracking of the respiratory motion. Section 3 presents a slight modification of the state-of-the-art 4DMRI sequence and two novel volume reconstruction techniques, both based on manifold learners. In addition, the construction of motion models from well-chosen data and the motion reconstruction from sparsely measured points is described. A one-dimensional breathing signal predictor based on kernel density estimation accomplishes Section 3. The quality of the reconstructions is evaluated by the prediction performance of the proposed methods and compared to the atlas-based prediction methods presented in Chapter 3 of this thesis. The paper closes with the discussion of the results.

Fast Manifold-based 4DMRI for Patient-Specific Organ Motion Modelling in MRgHIFU

Patrik Arnold, Silvio Pflugi, Frank Preiswerk, Beat Fasel, Philippe C. Cattin

Medical Image Analysis Center, University of Basel, 4000 Basel, Switzerland

E-mail: patrik.arnold@unibas.ch, philippe.cattin@unibas.ch

Abstract. MR-guided High Intensity Focused Ultrasound is an emerging non-invasive technique allowing to deposit sharply localised energy deep within the body, without affecting the surrounding tissues. This, however, implies exact knowledge of the target's position when treating mobile organs during the respiratory cycle. Existing motion modelling techniques either follow a population-based model approach, with the limitation that the model is organ, treatment-setup and patient-pose dependent and therefore not applicable to the variety of possible treatment scenarios. On the other side, subject-specific model approaches are considered, which can be very time consuming to acquire. Subject-specific models, in contrast, ensure customised treatment procedures providing comprehensive application. To address this issues, the minimal amount of required 3D data for the modelling process as well as novel 3D reconstruction techniques and minimalist model building approaches are presented. We conclude that the pre-operative data acquisition and processing time can be drastically reduced without significantly losing prediction accuracy.

1. Introduction

Respiratory-induced organ motion and its compensation are challenging tasks in the treatment of pathological tissue in mobile abdominal organs. Regardless of whether the tissue is treated by radiotherapy or Magnetic Resonance guided High Intensity Focused Ultrasound (MRgHIFU), accurate knowledge of the targets position is crucial for a safe and efficient therapy. Focused ultrasound, has the unique capability of depositing sharply localised energy deep within the body causing thermal ablation without damaging the surrounding tissue. Phased-array High Intensity Focused Ultrasound (HIFU) devices reach oblate ellipsoid lesion sizes of around 2 mm width and 4 mm length depending on the exposure time, acoustic power and sonication pattern [7]. Beam spot steering ranges up to 70 mm [3] and allow the continuous application of HIFU energy to the mobile target without interruption, which is essential to achieve the targeted thermal dose required for complete coagulation within typical exposure times of about 20s. Continuous alignment of the target and the therapy beam is only possible, if the method is able to predict the target's position at some future time in order to account for

the time delay between the acquisition of the current target position (surrogate marker) and the response of the system to adapt the treatment focus. To this end, the target has to be tracked in near real-time and an estimate of the target's future position has to be formed.

During an MRgHIFU treatment the patient is located within the Magnetic Resonance (MR) scanner enabling a completely non-invasive image-guided intervention. Although MR imaging provides excellent soft tissue contrast for target tracking, the MR scan-time is mainly required for the temperature feedback control of the HIFU system to determine the thermal dose delivered to a tumour. Therefore, extensive 3D or even 2D target tracking is impossible to achieve during treatment. Non MR-based external surrogate markers (*e.g.* optical chest wall trackers, breathing belt) or fast 1D MR navigator pencil beams [11] would thus be ideal for real-time tracking the respiratory-induced organ motion, demanding only for marginal MR scan-time. However, respiratory motion of the liver occurs in 3D space, whereby the main component is a cranio-caudal shift, usually in the range of 0.5-2.5 cm for quiet breathing [8, 12] and additionally the liver shows motion in anterior-posterior (1-12 mm) and left-right direction (1-3 mm) as well as non-rigid deformations (up to 2 cm) as quantified by [12]. Discarding the multidimensional target tracking demands for a technique to estimate the 3D target motion from a signal tracked in 1D. This estimation must be based on an underlying model.

In the literature, we can basically find three main types of such models, namely, subject-specific motion models [2, 9], population-based motion models [1, 10] and the combination of both [2]. Whilst the performance in terms of accuracy of subject-specific models is very high since the subject's motion itself is observed and learned, the acquisition and processing of 3D motion data is, still nowadays, very time consuming and hard to achieve in a clinical environment. Moreover, second order organ deformations occurring over long time scales, the so called drifts, can quickly invalidate the initially acquired model as shown in [1]. A short break between data acquisition for the model and the final treatment, hence, a fast model creation procedure is crucial for an appropriate usage of such a model. Using a population-based statistical model has the advantage of dispensing with the need of extensive pre-treatment volume imaging and data processing. On the other side, sophisticated and expert manual user input is required to establish correspondence among the patient and the motion model [10]. The central issue, however, is the limited applicability of population-based models. Possible variations in the treatment setup, such as additionally added Ultrasound devices to track the respiratory motion, changes in patient pose or the application to organs in an advanced stage of the disease, thus, evoking extraordinary shape, elasticity and motion pattern changes, can quickly invalidate the usage of such a population-based model. To this end, the creation of a patient-specific model in a clinically acceptable time would not only enable its application to other moving organs, but also to various treatment setups and multiple treatment sessions since such a model is independent of an organ specific database and fast to acquire.

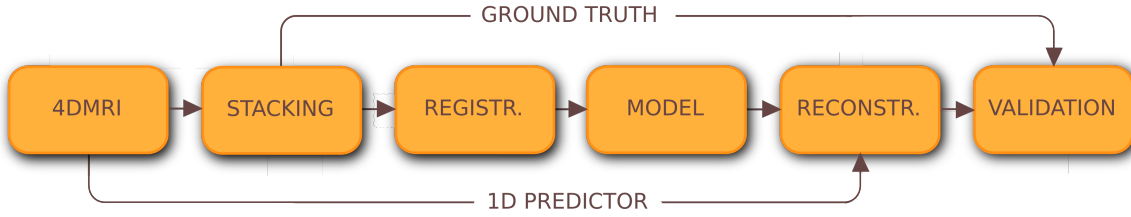


Figure 1. Schematic flowchart of the overall motion prediction procedure. Based on 4DMRI data, volumes are retrospectively reconstructed, registered and used to construct motion models given different sizes of datasets are built and evaluated on the ground truth data.

To address these issues, we propose time-optimised acquisition sequences combined with novel 3D reconstruction techniques and a minimalist model building approach. Figure 1 illustrates the overall prediction scheme consisting of data acquisition (4DMRI), data processing, in particular volume reconstruction and 3D registration, model building, temporal prediction combined with a spatial motion reconstruction and the evaluation of the prediction performance. The non-rigid volume to volume registration is next to the data acquisition the most time consuming step in the pre-operative model building process. Therefore, reducing the amount of registrations is an important step to speed up the overall procedure. The goal of this study is, thus, to investigate the minimal amount of required data to construct reasonable models for temporal as well as spatial motion prediction. Motivated by [6] and [20], manifold learners are used to resolve the respiratory state of the liver and to parametrise the breathing cycle. Based on regular 4DMRI data [18], two different *stacking* approaches are used for model building and evaluated given different lengths of acquisition times. With the respective model and an estimate of the temporal prediction, the displacement of the entire liver is reconstructed and evaluated on the ground truth data.

The organisation of the paper is as follows. In Section 2, the acquisition scheme and specifications of the clinical data to which the motion models will be applied is introduced. Section 3 describes the manifold-based 4DMRI approaches, the model building and motion reconstruction techniques as well as the 1D predictor used for respiratory signal prediction. Section 4 presents the performance evaluated on clinical data and Section 5 discusses the paper and summarises the results.

2. Material

2.1. Data Acquisition and Processing

The ground truth data for the experiments was acquired from 20 healthy volunteers in the course of 15-45 minutes acquisition sessions and is based on 4DMRI [18], a dynamic 2D MR imaging method capturing the respiratory motion during free breathing. The

acquisition results in two slice groups, the *data slices* covering the volume of interest consisting of around 30 slice positions covering the liver and a dedicated so-called *navigator slice* at a fixed position, both acquired in the sagittal plane with a voxel size of $1.4 \times 1.4 \times 4 \text{ mm}^3$ and a temporal resolution of 300-400 ms. An example of such an interleaved acquisition sequence with 2 *data slices* covering the volume of interest is illustrated in Figure 2(b). One complete volume pass with interleaved *navigator slices*, further called a *dynamic*, consists on an average of 60 slices, corresponding to around 12 seconds of MR acquisition time. All images were first processed using histogram matching in order to prevent the influence of different intensity values arising from tissue saturation effects, ensuring a similar distribution of intensity values among all images. In Sections 3.1 and 3.2, two retrospective stacking methods to reconstruct 3D volumes (*stacks*) are presented. *Standard* 4DMRI stacking [18], serving as ground truth and the novel *navigator-less* 4DMRI stacking. From the resulting *stacks*, dense spatio-temporal vector fields describing the motion between the different respiratory states of the liver are extracted. The first exhalation *stack* is manually segmented and taken as reference volume upon which the subsequent 3D *stacks* are incrementally registered from time-step to time-step by an affine, followed by a B-spline based non-rigid registration step [14], minimising the chance of getting trapped in local minima. For both, the initial rigid and the non-rigid registration, cross-correlation was chosen as similarity measure, whereby a control point resolution of 10 mm was used for the B-spline interpolation. The resulting vector fields describing the liver’s displacements relative to the reference volume are extracted at an isotropic grid with 15 mm resolution, corresponding to $N_s = 300\text{-}500$ grid points per subject s , depending on the size of the liver. These vector fields serve as the basic data for the motion models and their evaluation in cross-validation experiments.

2.2. Tracking of the Respiratory Motion

A key-concept of our approach is to estimate the 3D liver motion from a nearly real-time tracked 1-dimensional signal. As with any breathing-controlled tracking method, a system lag caused by target tracking, data processing and the time needed for refocusing the HIFU beam to the newly calculated target, has to be compensated for. Any signal correlated with respiration, such as the chest wall motion, could be taken here. In this research, the respective signal is extracted from the diaphragm motion of the *navigator slices* (see Figure 6(a)). Therefore, a manually selected template, located at the diaphragm, is tracked throughout all *navigator slices* by template matching with normalised cross-correlation as similarity measure. The resulting signal is linearly interpolated to obtain the intermediate respiratory states, where the ground truth 3D data for the evaluation is available [1]. The sampling rate of the *breathing signal* is, hence, 6-8 Hz, yielding look-ahead times of 125-166 ms. This signal is equivalent to the MR pencil beam navigator, as it is often acquired to obtain feedback information to

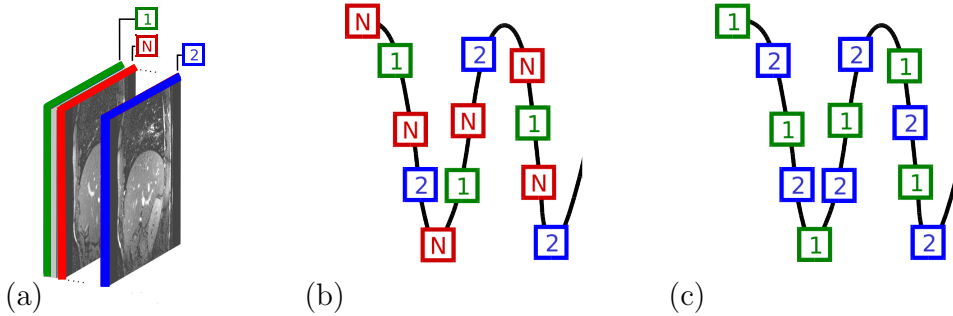


Figure 2. (a) Sagittal slices covering the volume of interest and dedicated *navigator* slice *N*. (b) Interleaved acquisition of *data slices* and *navigator slices* used for volume reconstruction in *Standard* 4DMRI and (c) acquisition scheme of *navigator-less* 4DMRI.

reposition the temperature mapping slice to resolve organ displacements [11] and used for the temporal prediction as described in Section 3.6.

3. Methods

In order to produce 3D volumes from the 2D images, *data slices* of all acquired slice positions, showing the same respiratory state, have to be found and stacked together, see Figure 2(a). The *standard* 4DMRI approach which is a slight modification of the one proposed in [18], is based on the interleaved acquisition of *data slices* and *navigator slices* *N* which are acquired immediately before and after the acquisition of the *data slices* as shown in Figure 2(b). The key-idea is, that two *data slices* show the same respiratory state of the liver, if their embracing *navigator slices* are similar. The novel proposed volume reconstruction approach is based on Locally Linear Embedding (LLE) [13], whereby the volumes are reconstructed without the need of the *navigator slices*, thus called *navigator-less* 4DMRI, resulting in the acquisition scheme shown in Figure 2(c).

3.1. Standard 4DMRI

As mentioned above, the *standard* 4DMRI approach is based on the similarity of the *navigator slices* which are acquired immediately before and after the *data slices*. To describe the similarity of the embracing slices, an image similarity measure based on Principal Component Analysis (PCA) [16] is used. In contrast to [18], where template matching of multiple manually defined regions are used for slice comparison, PCA does not need any manual user input. With this method, volumes for each subsequent pair of *navigator slices* are reconstructed, describing the motion for all observed time steps. The *navigator slices* of the *standard* 4DMRI sequence are given by the vectors $\mathbf{n} \in \mathbb{R}^{h \cdot w}$, where h and w are the height and width of the images, respectively. Let $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) \in \mathbb{R}^{h \cdot w \times m}$ be the data matrix containing m mean-free concatenated

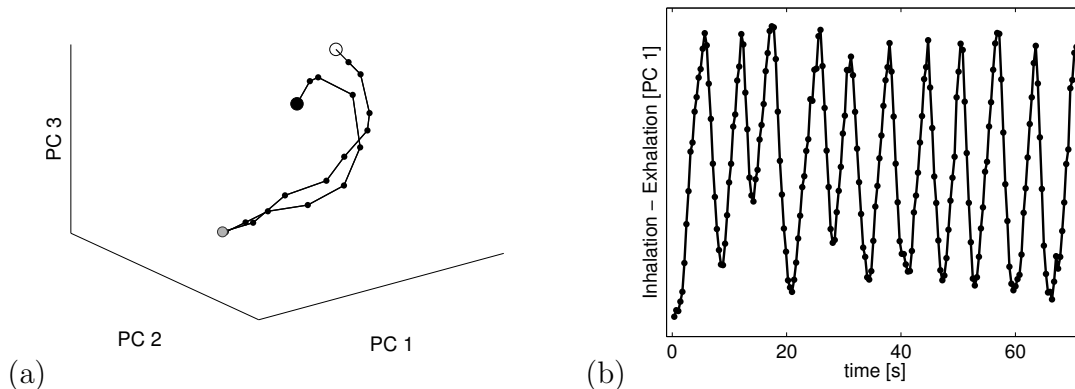


Figure 3. (a) Representation of navigator frames in 3D PCA subspace (from exhalation (●) over inhalation (◐) to exhalation (◑)). (b) Coefficients of the first PC resolving the respiratory state of the liver captured by the *navigator slices*.

navigator images $\mathbf{x}_k = \mathbf{n}_k - \mu$, with sample mean $\mu = \frac{1}{m} \sum_{k=1}^m \mathbf{n}_k$. In total, $m = 300$ images, 100 each from the beginning, the middle and the end of the acquisition are taken to derive the PCA. With the distributed selection of *navigator slices* anatomical changes occurring during the acquisition time are captured while the computational cost to derive the PCA is kept low due to the limited amount of selected input images. Applying PCA to \mathbf{X} yields the orthonormal matrix of principal components $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_{m-1})$ and their corresponding eigenvalues $\lambda_1, \dots, \lambda_{m-1}$ that give the standard deviation $\sigma_k = \sqrt{\lambda_k}$ of the principal components in descending order. The navigator images can be transformed into its coefficients \mathbf{c} , and thus mapped into the *navigator* space, by the simple operation:

$$\mathbf{c} = \mathbf{S}^T \cdot \text{diag}\left(\frac{1}{\sigma_k}\right) \cdot \mathbf{x}. \quad (1)$$

Hence, every *navigator* image can be described by its multidimensional coefficient vector \mathbf{c} . Figure 3(a) shows the mapping of one respiratory cycle into the *navigator* space represented by its first three PCA coefficients and Figure 3(b) shows 70s of the first principal component, nicely resolving the respiratory state of the organ. Based on these coefficients a complete 3D volume for each pair of embracing *navigator slices* can be reconstructed. For the reconstruction of the volumes the first eight principal components were considered to find the matches, whereof the two best matching slices were averaged for noise-reduction.

3.2. Navigator-less 4DMRI

The current state-of-the-art and above described *standard* 4DMRI approach uses the interleaved *navigator slices* to detect the respiratory state of the organ in order to retrospectively reconstruct 3D volumes. This approach is very well suited for long-time studies, where second order organ deformations, *i.e.* drifts, are the research objective [19]. Despite the robust reconstruction quality, the used acquisition scheme,

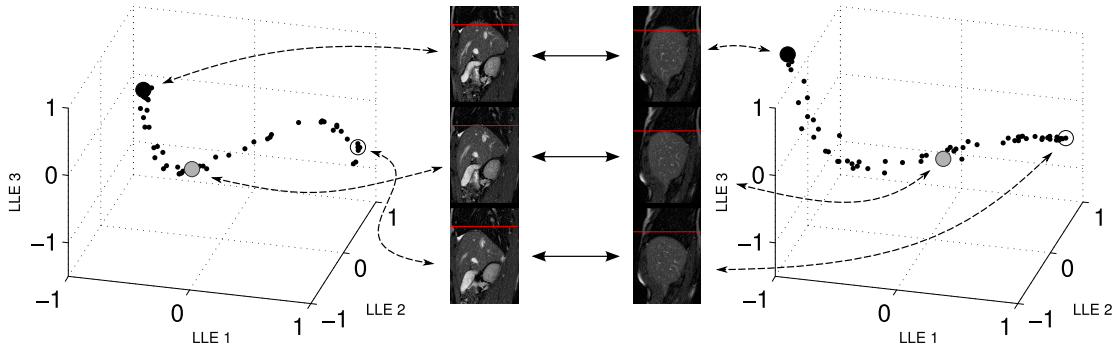


Figure 4. LLE slice matching: Representation of 60 *data slices* in 3D LLE subspace from inhalation (●) over mid-inhalation (◉) to exhalation (○) and corresponding *data slices* for two different slice positions of Subject 1. The red line indicates the mid-inhalation position of the diaphragm at the respective slice position.

however, unavoidably results in the measurement of actually redundant data, namely the extensive acquisition of the *navigator slices*. 50% of all acquired images are only used for retrospective volume stacking, but are actually of no further interest. The possibility of retrospectively reconstructing volumes without the need of interleaved *navigator slices* would not only end in 50% less acquisition time, but also double the frame acquisition rate. Using a manifold learner, it is possible to detect the given distribution of respiratory states per acquired slice position on the *data slices* themselves. Among many manifold techniques from the “Matlab Toolbox for Dimensionality Reduction” [17], Locally Linear Embedding (LLE) was found to work best for the purpose, meaning that LLE detects remarkably well the underlying structure of our data (see Figure 4). LLE is a non-linear unsupervised learning method that embeds the high dimensional data into a neighbourhood preserving lower dimensional space [13, 15]. Applying LLE to each *data slice* position, allows to compute a low dimensional embedding of our high dimensional input images ($125 \times 176 = 4400$ pixels) and maps every image into a lower dimensional space. The dimension of the subspace is chosen to be 3 which allows to visually examine the slice grouping process. We benefit from the property, that *data slices* showing the same respiratory state lie close to each other in the embedding space and, thus, allow to separate them. LLE is applied to all acquired slice positions individually yielding a subspace mapping for each slice position. The 3D representation depicted in Figure 4 nicely shows the ability of the method to reveal the respiratory motion contained in the images shown for two different slice positions. The embedded images form trajectories in the LLE-space describing the respiratory cycle from inhalation (●) over mid-inhalation (◉) to the exhalation position (○). The *data slices* represent the highlighted points on the trajectories. In order to find similar states among the different slice locations, the mapped data is parametrised by cubic splines. For cubic spline fitting, the order of the input data in terms of inhalation depth has to be known. Sorting by LLE-components is not robust enough to detect the sequence,

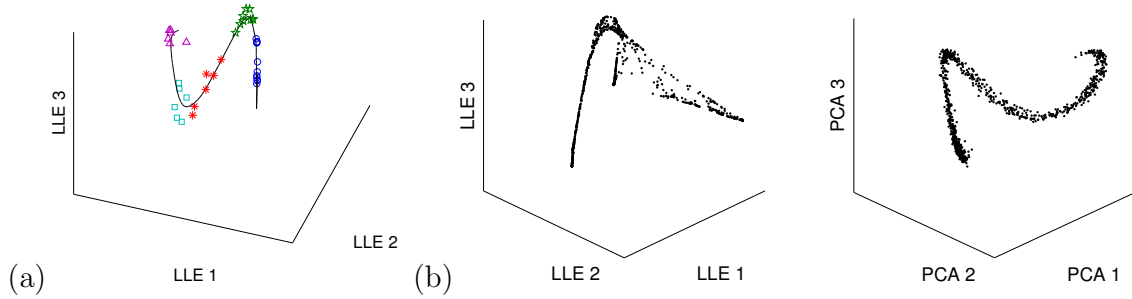


Figure 5. (a) LLE-coefficients in 3D space after grouping them around different states along the cubic spline fit. (b) Representation of LLE- and PCA-coefficients computed from 1000 *navigator slices* for subject 6 for a central slice position.

therefore template matching with cross-correlation is used to compare the diaphragm height of all images acquired at one position. This provides the order of the data for fitting and the number of diaphragm heights, *e.g.* the number of observed states for each slice position. The fitted cubic splines are subdivided into the number of observed respiratory states using arc-length discretisation whereby every data point representing an input image is assigned to the closest point on the spline. This clusters the data according their respiratory state and enables to find all acquired images, for example at the exhalation phase, of all slice positions, as visualised by different colours and markers in Figure 5(a). The data point, the image with minimal distance to the middle point of the corresponding spline piece, respectively, is chosen for 3D stacking. Again, the two best matches were averaged for noise-reduction. This results in one reconstructed *stack* per respiratory state which is used for model building. Note, that in contrast to standard 4DMRI, exhalation and inhalation phase are not distinguished.

3.3. LLE vs. PCA

The reason that PCA and not LLE is applied to the *navigator slices* in order to reconstruct the ground truth data is well illustrated by Figure 5. Figure 5(a) shows the 3D representation of 60 images and Figure 5(b) the mapping of 1000 *navigator* images in the respective embeddings. While LLE is better suited than PCA to accurately resolve the respiratory phase for short acquisition times containing up to 60 images as illustrated in Figure 4 and Figure 5(a), it fails for large sets of images. Due to the non-linearity of LLE, the mappings are very sensitive to small changes in the images, such as for example intestinal activities, which can occur over longer time scales, thus having a larger impact on the LLE-coefficients than the change of the respiratory phase. PCA on the other hand robustly resolves the respiratory state over long time scales as shown in Figure 5(b) (*right*). In general, we obtained better *navigator*-based reconstructions using PCA,

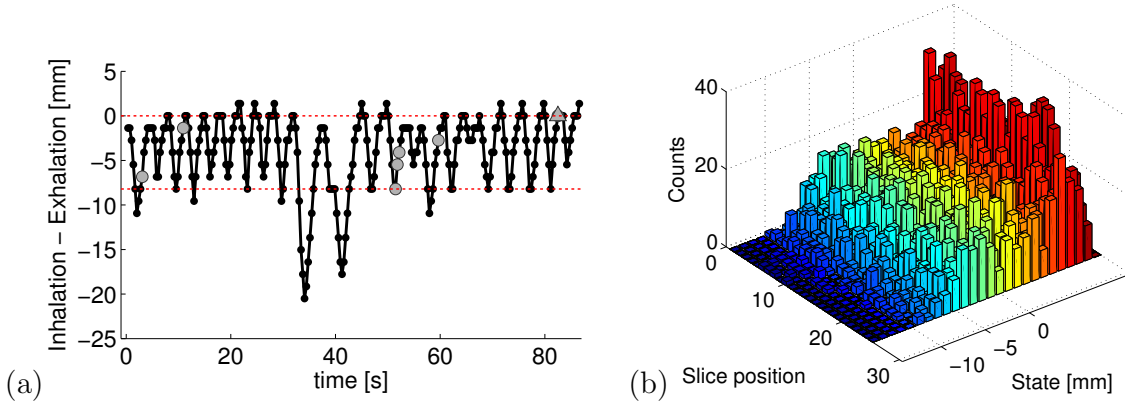


Figure 6. (a) Selection of best quality *stacks* \circ and exhalation *stack* \triangle between median exhalation and median inhalation position (---) used for model building and (b) the corresponding image distribution after 80 seconds of *standard* 4DMRI acquisition.

whereas on the other side LLE yielded better reconstructions with the *navigator-less* reconstruction technique.

3.4. Construction of Motion Models

To obtain reasonable models built from the least amount of data and thus reduce CPU-intensive and time-consuming data processing, only a handful of *stacks* are registered and used for modelling. Whilst for *navigator-less* 4DMRI only one *stack* per respiratory state is reconstructed, the number of *stacks* have to be reduced in the case of *standard* 4DMRI, picking out only the most representative *stacks* of the acquisition. Figure 6(a) illustrates this selection process for a *standard* 4DMRI sequence of around 80 seconds duration. All volumes are reconstructed for each time point (\bullet) and the quality of the *stack* is assessed by the sum of squared differences (SSD) between the neighbouring slices and averaged over the volume. From each respiratory state, the volume with minimal SSD (\circ) is taken and registered upon the exhalation *stack* (\triangle). Figure 6(a) shows the corresponding distribution of the *data slices* of the same scan with the bars illustrating the number of available *data slices* (counts) per given slice position and respiratory state. Obviously, not all bins are equally filled after 80 seconds of acquisition and not all respiratory states can be fully reconstructed. The maximum exhalation and inhalation states are determined by the median of all exhalation and inhalation position, respectively. All the *stacks* outside this range are assumed to be not a representative for the respective respiratory state as not all slice positions in the given state are available for stacking. These *stacks* are therefore not registered and used to build the model. This results in typically 7-10 reconstructed *stacks* per model, depending on the breathing depth of the respective subject. Deep inhalation states which are thus not included in the model can nevertheless be estimated by the prediction capabilities of the statistical model. Given a selection of *stacks* of one of the two 4DMRI acquisition methods, the *stacks* are registered and the resulting motion

fields are used to build the models. Let $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M)$ be the data matrix, where each column represents one of M observations, *i.e.* the organs displacement at a specific time point, respectively. The observations are described by a $3p$ -dimensional vector, containing the x, y, z -displacements of each of the p grid points within the liver: $\mathbf{v}_k = (\Delta x_1, \Delta y_1, \Delta z_1, \dots, \Delta x_p, \Delta y_p, \Delta z_p)'$. Equally, as described in Section 3.1, the original data is centered by $\mathbf{X} = \mathbf{V} - \bar{\mathbf{v}}$ with mean motion $\bar{\mathbf{v}} = \frac{1}{M} \sum_{k=1}^M \mathbf{v}_k$. Note, that \mathbf{X} does not contain any shape information, but only the relative displacements with respect to the manually segmented reference volume. Applying PCA to the data, the displacements of the organ can be described as linear combinations of the Eigenmotion vectors \mathbf{s}_k of $\mathbf{S} = (\mathbf{s}_1, \mathbf{s}_2, \dots)$ weighted by the coefficients \mathbf{c} , as given by Equation (1):

$$\mathbf{X} = \sum_{k=1}^{M-1} c_k \sigma_k \mathbf{s}_k \approx \sum_{k=1}^{M'} c_k \sigma_k \mathbf{s}_k. \quad (2)$$

Hereby, σ_k are the standard deviations within the data along each Eigenmotion s_k , sorted in a descending order. The organ motion can be approximated by a linear combination of the Eigenmotion corresponding to the largest M' Eigenvalues, resulting in a model that covers a certain amount of the data variance.

3.5. Reconstruction of Full Motion From Sparse Measurements

For many motion management applications it is not possible to measure the 3-dimensional position of the entire organ. Indeed this is the case for our MRgHIFU scenario, where surrogate markers such as the displacement \mathbf{d} of one point in 1D is measured. From the displacement \mathbf{d} , we wish to derive the entire organ displacement $\hat{\mathbf{v}}$. In order to estimate this complete displacement, further referred as to *reconstruction*, we use the framework of Bayesian estimation elaborated in [4]. The proposed technique allows *reconstructions* even though the displacement measure of only one single point in 1D is available. Basically, the most likely model coefficients $\hat{\mathbf{c}}$ for the full vector \mathbf{v} given an incomplete measurement $\mathbf{d} \in \mathbb{R}^l, l < 3p$ of \mathbf{v} has to be found. This can be done by minimising the following expression:

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

with $\mathbf{Q} = \mathbf{L}\mathbf{S} \cdot \text{diag}(\sigma_k)$, where \mathbf{L} represents a subspace mapping $\mathbf{L} : \mathbb{R}^{3p} \mapsto \mathbb{R}^l$. In the case of a noisy or incorrect assumption \mathbf{d} , tuning the regularisation factor η allows for reconstructions closer to the average quantified by the Mahalanobis distance $\|\mathbf{c}\|^2$. Solving Eq. 3 for $\hat{\mathbf{c}}$ with the Singular Value Decomposition of $\mathbf{Q} = \bar{\mathbf{V}}\mathbf{W}\mathbf{V}^T$, yields:

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

Using Eq. 4, the most probable organ displacement given an incomplete measurement \mathbf{d} is given by:

$$\hat{\mathbf{v}} = \mathbf{S} \cdot \text{diag}(\sigma_k) \hat{\mathbf{c}} + \bar{\mathbf{v}}. \quad (5)$$

The resulting reconstructions are compared to the ground truth data for the respective time step and averaged over the entire volume of the liver:

$$E_3 = \frac{1}{TN_s} \sum_{N_s} \sum^T \|\mathbf{v}_t - \hat{\mathbf{v}}_t\| . \quad (6)$$

3.6. Kernel Density Estimation for Respiratory Signal Prediction

For the treatment of mobile targets, such as it is the case for abdominal tumours, which are strongly affected by the respiratory breathing cycle, the prediction of respiratory motion is crucial for real-time target tracking. The target's future position must be estimated to compensate for system latencies, *i.e.* the latencies caused by signal acquisition, its processing and by readjusting the treatment beam to the newly calculated target.

To this end a Kernel Density Estimation (KDE) based predictor similar to the one in [5] is applied. Like all other predictors, KDE-prediction is based on training samples acquired at preceding time steps. The formulation is given for scalar observations acquired at uniform time intervals and prediction look-ahead times of an integer multiple of the sampling interval Δ_t . Suppose we are given a discrete signal s_i , acquired prior to the current time instant t , we can create the training samples $\chi_i = [s_{i-(h-1)}, s_{i-(h-2)}, \dots, s_i]$ with a signal history length h and corresponding target values $y_i = s_{i+1}$. Having a newly obtained measurement χ_t , the prediction \hat{y}_t is basically obtained through a weighted sum of the targets of the training samples:

$$\hat{y} = \frac{1}{\sum_i \omega_i} \sum_i \omega_i y_i . \quad (7)$$

The weights w_i determine how 'close' the test sample χ is to the training samples χ_i qualified by means of the Gaussian Kernel function K :

$$\omega_i = K(\chi_t, \chi_i) = \exp(-\alpha \|\chi_t - \chi_i\|^2), \quad (8)$$

with the kernel parameter $\alpha > 0$. The prediction is obtained by calculating the weights ω_i based on the training samples and inserting them into Equation 7. We implemented the moving window training, where only training samples close enough to the test sample are used for training. The window size should be chosen large enough to ensure reasonable learning. The parameters signal history length, training window size and kernel parameter α are optimised with respect to the prediction error $E_1 = \frac{1}{T} \sum^T \|y_t - \hat{y}_t\|$ in the training phase of duration T , that means during the acquisition of 3D data. Based on the prediction y_t , the 1-dimensional displacement \mathbf{d} is derived with respect to the model's exhalation *stack* and used for the reconstruction as described in Section 3.5.

4. Experiments and Results

The performance of the temporal prediction combined with the model-based motion reconstruction was evaluated on the ground truth data, which is based on all available

motion data from each of the 20 subjects. On average, displacements of the diaphragm from 5.5 mm to 15.2 mm in inferior-superior direction were observed. The proposed acquisition and stacking techniques are compared to the atlas-based motion prediction technique proposed by [2]. Figure 7 presents the visual comparison of the reconstruction

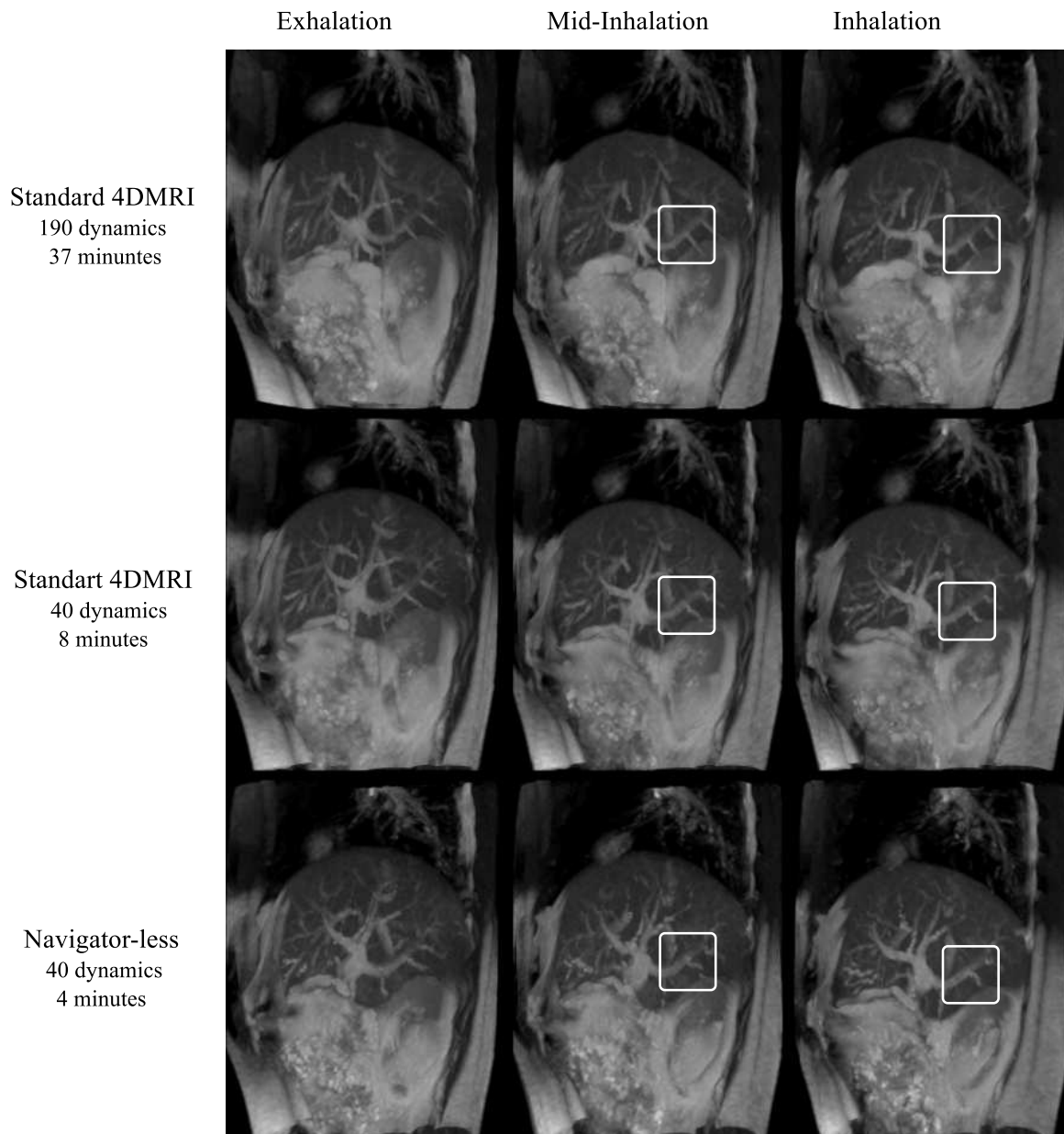


Figure 7. 3D reconstructions from exhalation over mid-inhalation to inhalation based on all available data, 40 *dynamics* of *standard* 4DMRI as well as *navigator-less* 4DMRI corresponding to 37 min, 8 min and 4 min of acquisition time, respectively.

quality of the used stacking approaches. The white boxes frame exemplary vessel structures used for quality assessment. Overall, the best quality is achieved for the PCA-based reconstructions, whereby the entire datasets were used and enough data to

reconstruct all states from inhalation to exhalation was available. The reconstruction based on shorter acquisition times, *i.e.* 8 minutes (40 *dynamics*), are of lower quality but still preserving the livers shape. Note, that for the LLE-method, theoretically, only half of the acquisition time (4 minutes) is required to achieve comparable quality. To find the relation between acquisition time and prediction performance, experiments with different sizes of training data have been performed. The above described procedure, *i.e.* stacking, registration and segmentation, was repeated for training sets of different lengths, namely 5, 10, 20, 30 and 40 *dynamics* corresponding to acquisition time ranging from 30s up to 8 min. For the *standard* 4DMRI sequence, the *navigator-less* 4DMRI and the atlas-based approach [2], a subject-specific model was generated for each of the training and datasets. The overall results are plotted in Figure 8(a) showing the prediction error averaged over all subjects depending on the respective acquisition time. *Navigator-less* 4DMRI performs best for short acquisition times below < 100 s. For longer acquisition times, on the other side, *standard* 4DMRI and atlas-based prediction yield smaller errors with increasing acquisition time. In particular, the experiments confirm an average prediction error below 2 mm after 60s of *standard* 4DMRI sequence and 30s *navigator-less* 4DMRI acquisition, respectively. The intended precision limit is set to 2 mm, thus, corresponding to the typical beam spot size of recent HIFU devices. As already shown in [2], secondary organ deformation, the so called drift, occurring over

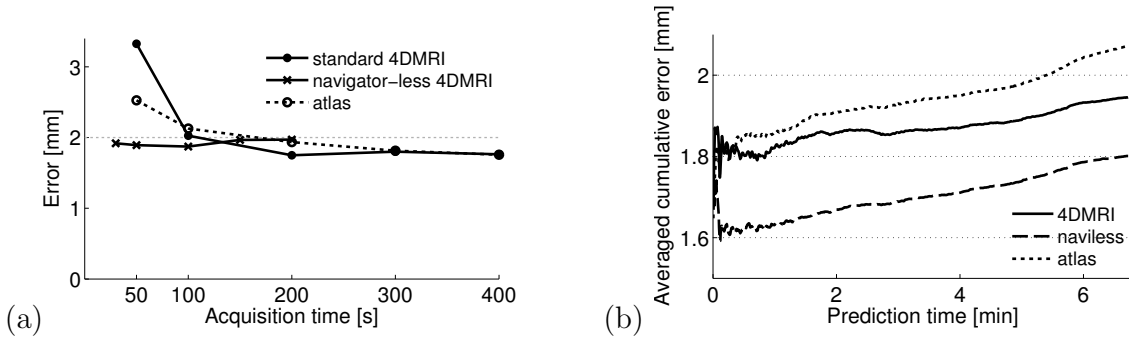


Figure 8. (a) Acquisition time versus prediction performance of *standard* 4DMRI, *navigator-less* 4DMRI and atlas-based prediction averaged over all 20 subjects.(b) Cumulative error averaged over all subject showing the impact of the organ’s drift for model predictions based on 100s of data acquisition.

longer time scales can quickly invalidate the applicability of an initially acquired subject-specific model. To quantify the drift-related error increase the cumulative average error over time was calculated over a period of about 6 min. Figure 8(b) demonstrates the impact of drift to prediction error for all evaluated methods, corresponding to an average growth rate of around 0.03 mm/min.

Table 1 presents the minimum of required acquisition time to achieve a motion prediction error below 2mm and the involved volume registrations to construct the respective model. The acquisition time and the number of non-rigid registrations are the most time consuming parts of the procedure. Therefore, the speedup factor in

respect to the atlas approach is shown. The time-saving in the case of the number of volume registrations is the speedup factor times the time per registration. Compared to the atlas-based prediction, the proposed methods show significant improvement in terms of speed. With *navigator-less* 4DMRI, an error below 2 mm is achieved after 30s of data acquisition and motion models constructed from on average only 8 volume registrations. This corresponds to a speedup factor of 6.6 and 32, respectively.

Table 1. Required acquisition time, number of involved volume registrations and respective speedup factor to achieve an overall motion prediction error below 2 mm.

Method	Acquisition		# Registration		Error
	Time	Speedup	Volumes	Speedup	
Atlas [2]	200 s	1	250	1	1.94 mm
Standard	100 s	2	9.9	26	1.99 mm
Navigator-less	30 s	6.6	8.2	32	1.87 mm

5. Discussion

Previous work on motion modelling [2] has shown that subject-specific motion models based on data acquired at a limited interval are suitable to predict the liver’s motion for a short time, before the organs drift invalidates the applicability of the model over time. If no population-based statistics or other prior knowledge about the specific organs long-term deformation is available to compensate for the occurring anatomical changes as for example described in [19], one has to account for larger errors with elapsed time. The time point of acquiring the model and the actual treatment should, thus, lie close together in order to ensure the models validity.

Within this context, the objective of this study was to investigate the impact of short acquisition times and minimalist model building techniques in terms of prediction performance. Subject-specific models have been created from test-sets of different duration reaching from 30-400s and their prediction capability have been evaluated against the ground truth data taken from 20 healthy volunteers. In particular, we could show in this study that accurate and robust tracking and prediction performance with minimal amount of data acquisition and processing can be achieved. The results presented in Figure 8(a) suggest to use *navigator-less* 4DMRI for short acquisition times up to 100s and *standard* 4DMRI otherwise. Due to the restriction of only taking *data slices* of the same respiratory phase for the volume reconstruction in the case of *standard* 4DMRI and the atlas-based approach, there is not enough data available to reconstruct reasonable *stacks* covering all breathing depths and thus, resulting in higher prediction errors for short acquisition times. For longer data acquisition times, on the other side, more exact *stacks* are obtained. For less than 30 seconds of acquisition, which corresponds to approximately 6 *data slices* per slice position, LLE fails to resolve an

appropriate mapping for all of the subjects. The error development over time shown in Figure 8(b) illustrates the impact of the organ's drift, quickly worsening the prediction accuracy of the models, which coincide with the findings in [2]. The growth rate of the error averaged over all subjects can be quantified to around 0.03 mm/min, which is negligible for short treatment times. Table 1 clearly points out a significant reduction of required volume registrations, yielding 26-32 times less computational cost for the proposed methods and, therefore relevantly reduces the delay between data acquisition and model application.

The proposed methods allow the fast acquisition of patient-specific 3D motion data and the efficient construction of patient-specific motion models applicable to every organ, target, patient and treatment setup, hence, overcoming the lack of generality of population-based models without sacrificing precision.

Acknowledgement

This work has been supported by the research network of the Swiss National Science Foundation-Project Nr. CR32I3_125499.

References

- [1] Patrik Arnold, Frank Preiswerk, Beat Fasel, Rares Salomir, Klaus Scheffler, and Philippe Cattin. Model-Based Respiratory Motion Compensation in MRgHIFU. *IPCAI*, LNAI 7330:54–63, 2012.
- [2] Patrik Arnold, Frank Preiswerk, Beat Fasel, Rares Salomir, Klaus Scheffler, and Philippe C. Cattin. 3D Organ Motion Prediction for MR-Guided High Intensity Focused Ultrasound. In *MICCAI*, volume LNCS 6892, pages 623–630, 2011.
- [3] Vincent Auboiroux, Erik Dumont, Lorena Petrusca, Magalie Viallon, and Rares Salomir. An MR-compliant phased-array HIFU transducer with augmented steering range, dedicated to abdominal thermotherapy. *Physics in Medicine and Biology*, 56(12):3563, 2011.
- [4] Volker Blanz and Thomas Vetter. Reconstructing the Complete 3D Shape of Faces from Partial Information. *Informationstechnik und Technische Informatik*, 44(6):295–302, 2002.
- [5] Ruan Dan. Kernel density estimation-based real-time prediction for respiratory motion. *PHYSICS IN MEDICINE AND BIOLOGY*, 55:1311–1326, 2010.
- [6] Manfred Georg, Richard Souvenir, Andrew Hope, and Robert Pless. Manifold learning for 4D CT reconstruction of the lung. In *IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*, 2008.
- [7] J W Hand, A Shaw, N Sadhoo, S Rajagopal, R J Dickinson, and L R Gavrilov. A random phased array device for delivery of high intensity focused ultrasound. *Physics in Medicine and Biology*, 54(19):5675, 2009.
- [8] Derek L.G. Hill, Colin Studholme, and David J. Hawkes. Voxel similarity measures for automated image registration. *SPIE*, 2359:205–16, 1994.
- [9] Ruijiang Li, John H. Lewis, Xun Jia, Tianyu Zhao, Weifeng Liu, Sara Wuenschel, James Lamb, Deshan Yang, Daniel A Low, and Steve B Jiang. On a pca-based lung motion model. *PHYSICS IN MEDICINE AND BIOLOGY*, 56:6009–6030, 2011.
- [10] Frank Preiswerk, Patrik Arnold, Beat Fasel, and Philippe C. Cattin. A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements. *Springer LNCS*, 7029:207–214, 2011.
- [11] M. Ries, B. D. Senneville, S. Roujol, Y. Berber, B. Quesson, and C. Moonen. Real-time 3d target

- tracking in mri guided focused ultrasound ablations in moving tissues. *Magnetic Resonance in Medicine*, 64:01704–1712, 2010.
- [12] T Rohlfing, CR Maurer, WG O’Dell, and JH Zhong. Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated MR images. *MEDICAL PHYSICS*, 31(3):427–432, March 2004.
- [13] Sam Roweis and Lawrence Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000.
- [14] 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. *Transactions on Medical Imaging*, 18:712–21, 1999.
- [15] Lawrence K. Saul and Sam T. Roweis. Think globally, fit locally: Unsupervised learning of low dimensional manifolds. *Journal of Machine Learning Research*, 4:119–155, 2003.
- [16] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. *IEEE*, pages 586–591, 1991.
- [17] Laurens van der Maaten. Matlab toolbox for dimensionality reduction. v0.7.2b, 2012.
- [18] Martin von Siebenthal, Gabor Székely, Urs Gamper, Peter Boesiger, Antony Lomax, and Ph. Cattin. 4D MR imaging of respiratory organ motion and its variability. *Physics in Medicine and Biology*, 52:1547–64, 2007.
- [19] Martin von Siebenthal, Gabor Székely, Antony Lomax, and Philippe Cattin. Inter-subject modelling of liver deformation during radiation therapy. In *MICCAI*, LNCS 4791, pages 659–66, 2007.
- [20] Mehmet Yigitsoy, Christian Wachinger, and Nassir Navab. Manifold learning for image-based breathing gating in mri. *SPIE*, 7962:796210–1, 2011.

6 Discussion and Conclusion

- Contents** Our investigations in the previous chapters have led to many observations and new insights. Three approaches to manage respiratory-induced organ motion have been presented. Each method has its respective advantages and disadvantages. This chapter discusses the main observations and provides answers to the two research questions of the thesis as posed in Chapter 1. Finally, directions for future research are presented.
- Outline** In Section 6.1 , we answer the two research questions based on the obtained results from the former chapters. Directions for future research are presented in Section 6.2.

6.1 Answers to the Research Questions

Several approaches to manage respiratory-induced organ motion have been presented in this thesis. This section addresses the research questions stated in Chapter 1 based on the observations made in the three previous chapters.

Research question 1: *How accurate can the 3D position of a target be predicted, given that the respiratory state of the liver can be tracked in 1D?*

The target prediction methods proposed in this thesis consist of two parts. First, the temporal prediction of the breathing signal and second, the spatial motion prediction of the liver based on this breath prediction.

The temporal prediction of the respiratory state was either realised by a simple pattern matching approach as described in Chapters 3 and 4 or by a more sophisticated predictor based on Kernel Density Estimation (KDE) as described in Chapter 5. Both approaches are able to robustly predict the future respiratory state of the liver with a prediction error of approximately 0.5 mm. The advantage of the KDE-based predictor is that it is able to produce continuous predictions. The pattern matching based predictor, in contrast, can only predict patterns that have been observed before through template matching.

No matter which temporal predictor is applied, the resulting estimate of the future respiratory state is used to reconstruct the motion of the entire liver. This reconstruction is based on an underlying motion model. In order to quantify the 3D motion prediction accuracy, we investigated the performance of subject-specific models combined with a population-based drift model in Chapter 3 (atlas + drift), pure population-based motion models in Chapter 4 and pure subject-specific motion models as presented in Chapter 5. In Table 6.1 the results of our investigations are summarised. The best results in terms

Table 6.1: Summary of prediction results achieved during this study.

Method	Error [mm]
Atlas + Drift Model (Chapter 3)	1.1 ± 0.4
Population-based Model (Chapter 4)	1.7 ± 0.5
Subject-specific Model (Chapter 5)	1.9 ± 0.6

of accuracy are achieved with the combination of a subject-specific motion model and a population-based drift model (Chapter 3). In this case, the subject-specific motion was observed up to 13 minutes and stored in an atlas. Based on this atlas, the prediction

performance was evaluated for the subsequent 13 minutes. It was shown that the drift can quickly invalidate the initially acquired atlas. Therefore, the prediction performance could be significantly improved by incorporating population-based drift information.

To avoid extensive patient-specific data acquisition and processing as it was done in Chapter 3, a pure population-based motion modelling approach was investigated next in Chapter 4. Although no patient-specific data except the exhalation shape and position was known, the prediction performance of the model is high and robust. As the experiments were evaluated for a shorter period, *i.e.* 7-11 minutes, no drift correction was applied.

Despite the good performance, population-based models are limited in their applicability. Slight changes in the treatment setup, as for example a change in the patient pose, can quickly invalidate population-based motion models. Therefore, the research focus moved towards the fast acquisition of subject-specific motion models as presented in Chapter 5. These models are *per se* generally applicable since the model is acquired immediately before the treatment and thus capturing all the setup and patient specific characteristics.

As the liver can significantly drift over time, initially acquired motion models can loose accuracy over time as shown in Chapter 3. Therefore, a short time gap between model acquisition and treatment start is crucial for accurate model-based motion prediction. These thoughts brought us to the second research question:

Research question 2: *How much pre-operative data is required to construct models suitable for motion prediction?*

In order to answer this question, subject-specific motion models based on different lengths of 4DMRI datasets were constructed and evaluated. It could be shown that short 4DMRI acquisition times in the range of 30-200 seconds are sufficient to construct motion models with an average prediction performance below 2 mm. The results are presented in detail in Chapter 5, Table 1. Although the reconstruction quality suffers from short 4DMRI acquisition times, the prediction performance of the resulting motion models is surprisingly good. The ability to perform accurate motion predictions despite the limited time for data acquisition is an important insight for the clinical relevance of the proposed methods.

This thesis presents possible motion compensation techniques for the non-invasive and radiation-free treatment of pathological tissue in moving abdominal organs in MRgHIFU. We conclude that patient-specific models are best suited to cover the large variety of possible MRgHIFU treatment scenarios.

6.2 Future Research

The future research focus should lie on the improvement of the currently developed techniques. So far, the prediction performance of the proposed methods have been validated on 4DMRI data through theoretical simulations. The next steps are therefore the integration of the technique in a real MRgHIFU system and its validation in vivo. To achieve a realisation of the proposed techniques in a clinical environment, we recommend the following improvements:

- Ensure a 4DMRI acquisition where all respiratory bins are more equally occupied. One way to achieve this is by using breath coaching. With the patient following a visual or audio signal, a more regular breathing pattern would be achieved. This benefits a more consistent data distribution and prevents very deep breaths.
- Alternatively, an adaptive 4DMRI slice acquisition sequence could be applied. In addition to the 4DMRI sequence, a pencil beam tracking the respiratory state can be acquired. By means of this pencil beam, a lookup table, containing information about the already captured *data slices* for a given breathing depth, could be created. In addition, the pencil beam would resolve the current respiratory state. Based on the lookup table and the measurement of the current state, the missing slices for the respective breathing depth could be captured.
- In case of empty bins during 3D volume reconstruction, the missing slices can be obtained through slice interpolation. This, however, can only be done for slices where both embracing slices are available.
- A short time gap between data acquisition and start of the treatment is crucial for accurate motion prediction. Therefore, a fast implementation of a non-rigid registration algorithm would significantly decrease the acquisition time.

Curriculum Vitae

PERSONAL DATA

Name **Patrik Arnold**
Date of Birth 02.02.1983
Place of Birth Bern, Switzerland

EDUCATION

2009 – 2012 **PhD in Biomedical Engineering**
Medical Image Analysis Center, Universität Basel
3D Organ Motion Compensation for MRgHIFU
Doktorvater: Prof. Dr. Philippe C. Cattin

2003 – 2008 **Master of Science in Physics**
Institut für Angewandte Physik, Universität Bern

REFERENCES

Prof. Dr. Philippe C. Cattin
Medical Image Analysis Center
Universität Basel
Spitalstrasse 21, Basel
+41 (0)61 265 96 55
philippe.cattin@unibas.ch

Dr. A. Murk
Institut für Angewandte Physik,
Universität Bern
Sidlerstrasse 5, 3012 Bern
+41 (0)31 631 89 11
axel.murk@mw.unibe.ch

Basel, December 4, 2012

List of Publications

The investigations performed during my PhD research resulted in the following publications and talks.

Papers

- **Patrik Arnold**, Frank Preiswerk, Beat Fasel, Rares Salomir, Klaus Scheffler, and Philippe Cattin. Model-Based Respiratory Motion Compensation in MRgHIFU. IPCAI, LNAI 7330:54–63, 2012.
- **Patrik Arnold**, Frank Preiswerk, Beat Fasel, Rares Salomir, Klaus Scheffler, and Philippe C. Cattin. 3D Organ Motion Prediction for MR-Guided High Intensity Focused Ultrasound. In MICCAI, volume LNCS 6892, pages 623–630, 2011
- **Patrik Arnold**, Silvio Pflugi, Frank Preiswerk, Beat Fasel and Philippe C. Cattin. Fast Manifold-based 4DMRI for Patient-Specific Organ Motion Modelling in MRgHIFU. *to be published*
- Frank Preiswerk, **Patrik Arnold**, Beat Fasel, and Philippe C. Cattin. Robust Tumour Tracking From 2D Imaging Using a Population-Based Statistical Motion Model. In IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA), 2012.
- Frank Preiswerk, **Patrik Arnold**, Beat Fasel, and Philippe C. Cattin. Towards More Precise, Minimally-Invasive Tumour Treatment Under Free Breathing. In 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE EMBC, 2012.
- Frank Preiswerk, **Patrik Arnold**, Beat Fasel, and Philippe C. Cattin. A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements. Springer LNCS, 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.
- Lorena Petrusca, Valeria De Luca, **Patrik Arnold**, Zarko Celicanin, Thomas Goget, Vincent Auboiroux, Magalie Viallon, Francesco Santini, Christine Tanner, Sylvain Terraz, Philippe Cattin, and Rares Salomir. Ultrasound/MR Hybrid Imaging: Truly Simultaneous Motion Monitoring in the Abdomen and Image Co-Registration. 2011.
- Lorena Petrusca, **Patrik Arnold**, Thomas Goget, Zarko Celicanin, Vincent Auboiroux, Magalie Viallon, Francesco Santini, Valeria De Luca, Sylvain Terraz, Christine Tanner, Klaus Scheffler, Christoph D. Becker, Philippe Cattin, and Rares Salomir. Simultaneous Ultrasound/MRI Motion Monitoring in the Abdomen. In Interventional MRI Symposium (iMRI), 2010.

Talks

- **MICCAI 2011**, Toronto, Canada
- Invited talk at **Accuray**, Headquarters, Sunnyvale, CA ,USA